СОВРЕМЕННЫЕ КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ В МОДУЛЬНЫХ МНОГОПРОЦЕССОРНЫХ СИСТЕМАХ:

ОПЫТ ИСПОЛЬЗОВАНИЯ, ИССЛЕДОВАНИЕ ОЦЕНОК ЭФФЕКТИВНОСТИ, ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ

МОНОГРАФИЯ

Днепропетровск, 2012

УДК 004
ББК

В.П. Иващенко
Е.А. Башков
Г.Г. Швачич
М.А.Ткач

Рецензенты:

доктор физико-математических наук, профессор ***В.Н. Моисеенко***
доктор технических наук, с.н.с. ***Ю.К. Тараненко***

Современные коммуникационные технологии в модульных многопроцессорных системах: опыт использования, исследование оценок эффективности, перспективы применения : монография / В.П. Иващенко, Е.А. Башков, Г.Г. Швачич, М.А.Ткач. – Днепропетровск, 2012. – 139 с.

В монографии показаны пути повышения эффективности многопроцессорных кластерных систем за счет реорганизации архитектуры сетевого интерфейса. Проведен анализ исследования перспектив применения современных коммуникационных технологий в многопроцессорных кластерных системах. Основное внимание уделяется особенностям влияния сетевого интерфейса многопроцессорной вычислительной системы на оценки ее эффективности. Раскрыты особенности сопряжения модульных многопроцессорных систем для расширения вычислительных возможностей при решении сильносвязанных задач.

Для научных сотрудников и аспирантов, специализирующихся в области конструирования многопроцессорных вычислительных систем, всех, кто серьезно интересуется проблемами влияния коммуникационных технологий на эффективность распараллеливания в многопроцессорных системах.

СОДЕРЖАНИЕ

Раздел 1.
ИССЛЕДОВАНИЕ ВЛИЯНИЯ СЕТЕВОГО ИНТЕРФЕЙСА НА ЭФФЕКТИВНОСТЬ МОДУЛЬНОЙ МНОГОПРОЦЕССОРНОЙ КЛАСТЕРНОЙ СИСТЕМЫ

Введение

Применение параллельных вычислительных систем 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.

В данной работе рассматриваются так называемые “блейд” серверные решения многопроцессорных систем [1]. На основе сетевой технологии *Ethernet* было реализовано “блейд” серверное решение многопроцессорной системы, при которой 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 была сконструирована такого рода кластерная вычислительная система. Основные особенности конструирования ее архитектуры изложены в [2]. Система содержит отдельную реконфигурируемую сеть для обмена данных между вычислительными узлами, дополнительные управляемые коммутаторы, которые работают параллельно, промежуточные буферы памяти коммутаторов, а также предусматривает сетевую загрузку узлов и механизм резервирования ключевых компонентов.

Практически одновременно с появлением первых многопроцессорных систем возникла необходимость в оценке их эффективности, производительности, быстродействия и в последующем сравнении подобных вычислительных систем, учитывая эти параметры. 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. Указанные вопросы являются определяющими для процедуры моделирования широкого класса задач при помощи модульных многопроцессорных кластерных систем. В работе [3] достаточно детально изложены вопросы исследования эффективности такой многопроцессорной вычислительной системы.

Исследования, посвященные влияния сетевого интерфейса на эффективность модульных многопроцессорных вычислительных систем, в настоящее время приобретают важнейшее значение. Следовательно, определяющим фактором эффективности функционирования модульных многопроцессорных систем являются вопросы выбора, конструирования и организации их сетевых интерфейсов. С этих позиций и будет оцениваться эффективность многопроцессорной кластерной системы, которая исследуется в данной работе.

Ныне рынок сетевых технологий продолжает стабильно развиваться, хотя специалисты утверждают, что темпы его развития несколько меньше, чем отрасли ИТ в целом. Например, оптоволоконная связь имеет громадные резервы пропускной способности: частота несущих колебаний на несколько порядков превышает освоенные частоты модулирующего сигнала (например, у *Gigabit Ethernet*). Однако для использования этих резервов требуется дальнейшее развитие микроэлектроники, пока что позволяющей уверенно использовать скорости передачи до 10 Гбит/с в одном канале. Число оптических волокон в кабелях обычно составляет от 4 до 216. Срок службы волоконно-оптических кабелей: как правило, не менее 25 лет [4]. Это можно объяснить тем, что кабельная система является наиболее консервативным компонентом в инфраструктуре сетевых технологий. При этом срок службы кабельных систем, в среднем, составляет до десяти лет, а почти все остальные производители дают на свою продукцию гарантию не менее двадцати.

В настоящее время существует целый ряд различных высокоскоростных сетевых и коммуникационных аппаратных технологий, которые используются для связи вычислительных узлов при построении многопроцессорных кластерных систем. К наиболее распространенным можно отнести следующие: *Fast Ethertnet, Gigabit Ethernet, Myrinet, cLAN (Giganet), SCI, QsNetII (QSW), MEMORY CHANNEL, ServerNet II, InfiniBand, Flat Neighborhood* и др. Однако при выборе оптимального решения для построения персональных вычислительных кластеров (ПВК) часто отдают предпочтение *Gigabit Ethernet* [2, 5, 6]. Главная причина, позволяющая стандарту *Ethernet* претендовать на роль глобального решения при конструировании ПВК – это низкая стоимость соответствующих аппаратно-программных средств. Расходы на построение сети *Ethernet* составляют лишь пятую часть от того, что придется потратить на проектирование конкурирующей с ней, например, сети *SONET* [2].

В настоящее время можно ознакомиться с целым рядом исследований, посвященных тестированию различных сетевых технологий [7, 8], однако, наряду с этим, не получили должного развития или вообще отсутствуют работы, посвященные исследованию влияния различного типа сетевого интерфейса на эффективность модульных многопроцессорных вычислительных систем. При этом определяющим фактором эффективности функционирования модульных многопроцессорных систем являются вопросы выбора, конструирования и организации их сетевых интерфейсов.

Данная работа посвящена проблеме исследования оценок эффективности многопроцессорной вычислительной системы. Основное внимание уделяется влиянию сетевого интерфейса на оценки эффективности кластерной системы. Выведены аналитические соотношения основных оценок эффективности вычислений через параметры многопроцессорной системы. Кроме того, в работе рассмотрены перспективы применения современных коммуникационных технологий в многопроцессорных кластерных системах.

1.1. Объект исследований

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 логически упомянутая система *MPP* мало отличается от обычной локальной сети.

В данной работе в качестве объекта исследований рассматриваются так называемые “блейд” серверные решения многопроцессорных систем.

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.

Блок-схема исследуемой многопроцессорной вычислительной системы представлена на рис. 1.1. 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. 1.1) заключается в том, что все его вычислительные узлы *PM*001, *PN*001- *PN*005 содержат процессор (1) *C7 CPU*, присоединенный шиной *FSB* (*Front Side Bus* 533/400 МГц) к южному мосту *CN*700 (2) с интегрированным

видеоконтроллером *VIA UniChrome Pro* и видеовыходами *SVGA* (3), *TV* (4) и интерфейсом *AGP* 8*X* (5), а южный мост подключен к локальной памяти (6), которая работает по стандартам *DDR*2 533/400 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 *VT*8237A (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), поддерживающий интерфейсы *SATA*, *PATA* и режим *RAID*, а также присоединена шина *PCI Bus* с двумя разъемными соединениями *PCI* (10,11), в которых установлены сетевые интерфейсы, поддерживающие режимы *channel bonding* и *Gigabit Ethernet* (*Gi*_00*x*,1 12, *Gi*_00*x*,1 13). К мосту *VIA VT*8237*A* подключен интегрированный аудиоконтроллер (14) *VIA Viny*l™ *HD Audio*, контроллер клавиатуры и манипулятора мыши *PS*/2 (15), а также восемь высокоскоростных портов стандарта *USB* 2.0 (16) и контроллер *VT*1211 (17), который представляет собой полнофункциональный *Super I/O*-чип с контроллером дисковода гибких дисков, интерфейсом параллельного порта *IEEE*-1284, двумя последовательными портами 16*C*550-*UART*, контроллером *VFIR* (скоростной инфракрасный порт), игровым портом, который поддерживается двумя джойстиками, *MIDI*-интерфейсом и интерфейсом 4*M FLASH-ROM BIOS* (18), интегрированным сетевым интерфейсом, который может поддерживать режим сетевой загрузки и *Fast Ethernet P*M001.*i*01 (19).

Основой такого кластера являются, установленные 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, решающие 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, насколько это, возможно, является плотной и имеет необходимые условия для эффективного соединения компонентов. Главные узлы, узлы управления и узлы запоминающих устройств обеспечивают функции управления кластером – загрузка, управление устройствами, внешний ввод – вывод и т. д.

Персональный вычислительный кластер, блок-схема которого изображена на рис. 1.1, имеет такие размеры: ширина 19, высота 10,9, глубина 9 дюймов. Вес кластера составляет приблизительно 8 кг.

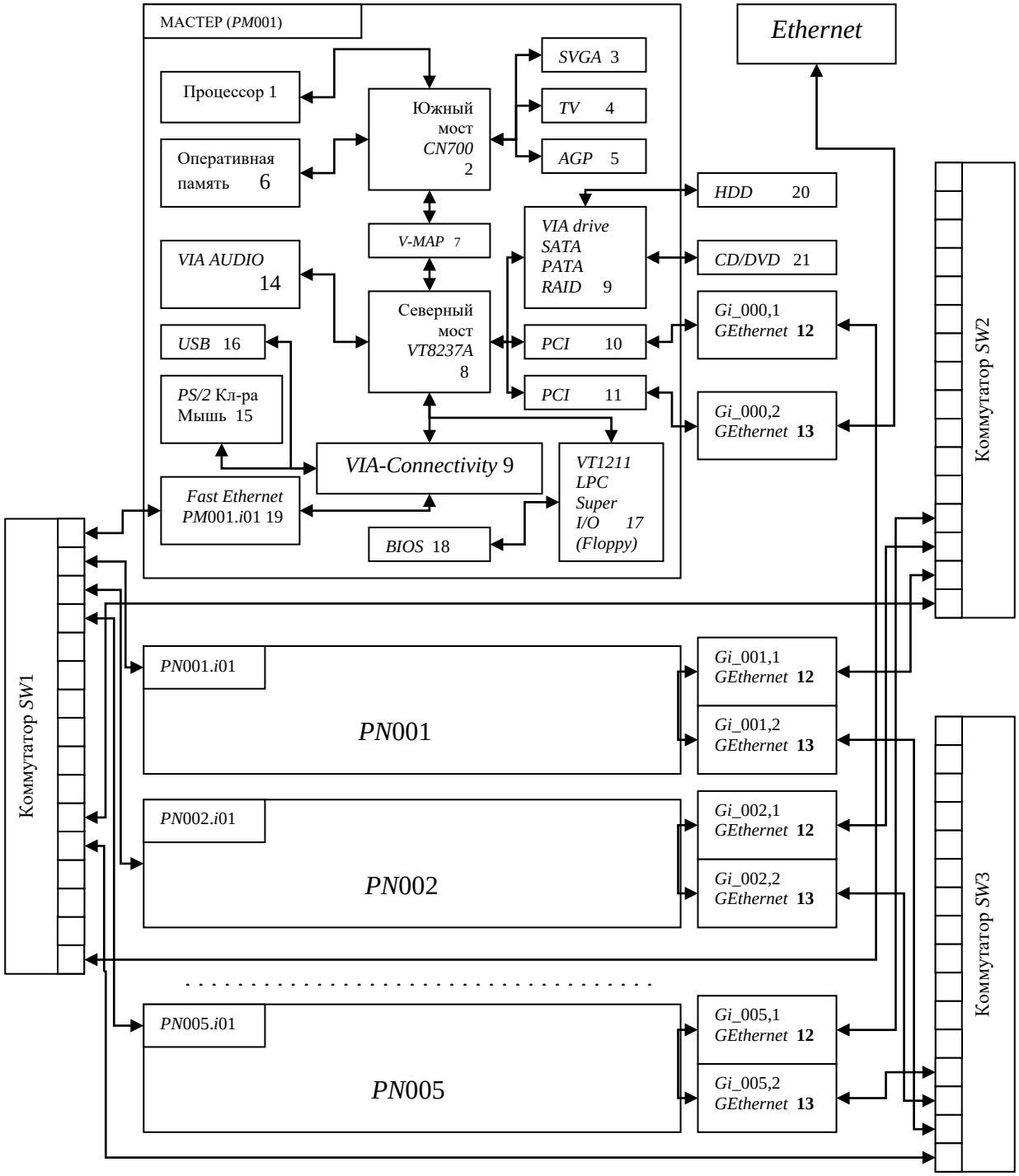

Рис. 1.1. Блок-схема модуля многопроцессорной системы

1.2. Постановка проблемы исследований, ее актуальность

В настоящее время появилась уникальная возможность создавать недорогие установки суперкомпьютерного типа – многопроцессорные вычислительные кластеры. До недавнего времени было сомнение в перспективности такого направления. Однако, при всех “за” и “против”, постоянным обитателям списка *Top*500 [9] – компаниям *Cray, Sun, Hewlett-Packard* и другим пришлось потесниться, пропустив вперед ряд кластерных решений. С другой стороны, в настоящее время рынок сетевых технологий интенсивно развивается, и производители сетевых решений на базе *cLAN, Myrinet, ServerNet, SCI* продолжают и дальше совершенствовать свои технологии, давая возможность практически без особых финансовых затрат конструировать свой собственный вариант суперкомпьютера.

Очевидно, что на сегодняшний день существует много различных вариантов построения кластерных вычислительных систем. Однако одно из основных различий в их конструировании лежит в области используемой сетевой технологии, выбор которой определяется, прежде всего, классом решаемых задач.

Например, в задачах металлургии при математическом моделировании скоростных режимов термической обработки длинномерных изделий одна из основных проблем может быть сформулирована так: имеем разностную сетку размерности M, время вычисления задачи, которая решается с использованием однопроцессорной системы, определяется величиной t. Этот параметр является критичным. Необходимо существенно уменьшить время вычислений, сохраняя значение M. В этой связи, вопросам эффективности, быстродействия и производительности уделяется основное внимание при конструировании кластерных систем. Итак, рассматривается задача уменьшения времени расчетов путем увеличения числа узлов кластерной системы. Такой подход ориентирован, например, на разработку новых технологических процессов (когда время вычислений являет собой критическую величину) [10, 11]. Кроме

того, аналогичные задачи часто приходится решать в медицине, военной технике и др.

Таким образом, тема конструирования кластерных многопроцессорных систем на сегодняшний день является актуальной, интересной и переживает этап своего бурного развития.

Ясно и другое, что при помощи высокопроизводительных кластеров найден эффективный способ решения широкого класса актуальных задач.

По нашему мнению, новый качественный этап развития многопроцессорных кластерных систем лежит в области использования новых современных сетевых технологий. Это объясняется следующим образом. Сеть кластерной вычислительной системы принципиально отличается от сети рабочих станций, хотя для построения кластера необходимы обычные сетевые карты и хабы/коммутаторы, которые применяются при организации сети рабочих станций. Однако в случае кластерной вычислительной системы имеется одна принципиальная особенность. *Сеть кластера, в первую очередь, предназначена не для связи компьютеров, а для связи вычислительных процессов*. В этой связи, чем выше будет пропускная способность вычислительной сети кластера, тем быстрее будут считаться пользовательские параллельные задачи, выполняемые на кластере. Таким образом, технические характеристики вычислительной сети приобретают первостепенное значение для многопроцессорных кластерных систем.

В настоящее время проблема выбора и анализа сетевых технологий для модульных многопроцессорных кластерных систем не получила должного развития. Кроме того, практически отсутствуют работы, посвященные исследованию влияния сетевых технологий на эффективность распараллеливания в модульных многопроцессорных кластерных системах. В этой связи, рассматриваемые в данной работе исследования, являются актуальными и, несомненно, вызовут интерес у соответствующих специалистов.

1.3. Цель и задачи исследований

Основная цель исследований, представленных в данной работе, заключается в дальнейшем развитии подхода, связанного с анализом оценок эффективности модульной многопроцессорной кластерной вычислительной системы. При этом основное внимание уделяется особенностям влияния сетевого интерфейса такой системы на оценки ее эффективности.

В результате проведенных исследований необходимо решить следующие задачи:

− выявить и установить основные режимы работы сетевого интерфейса в многопроцессорных кластерных системах;

− провести анализ основных режимов работы сетевого интерфейса в многопроцессорных кластерных системах и выявить их влияние на оценки эффективности распараллеливания;

− выявить пути повышения эффективности многопроцессорной кластерной системы за счет организации архитектуры ее сетевого интерфейса;

− вывести аналитические соотношения для определения оптимального числа узлов кластерной системы для различных режимов ее функционирования;

− для удобства определения оценок эффективности кластерной вычислительной многопроцессорной системы вывести основные аналитические соотношения через параметры исследуемой системы.

Кроме того, необходимо провести анализ перспектив применения современных коммуникационных технологий в многопроцессорных кластерных системах. Для этой цели необходимо решить следующие задачи:

− исследовать особенности формирования архитектуры сетевого интерфейса кластерной системы на основе применения технологий: *Myrinet, Fibre Channel, 10Gb Ethernet;*

− провести анализ согласования выбранного сетевого оборудования на основе моделирования основных сетевых коэффициентов кластерной системы;

− выполнить сравнительный анализ оценок эффективности многопроцессорной кластерной системы для различного типа сетевых технологий;

− для увеличения пропускной способности сети кластера применить технология "связывания каналов" (*channel bonding*);

− определить характер улучшения оценок эффективности модульной кластерной системы за счет реализации технологии *channel bonding*;

− определить влияние ценового фактора современных коммуникационных технологий на эффективность распараллеливания.

Качественный этап развития многопроцессорных кластерных систем лежит в области сопряжения нескольких модулей многопроцессорных систем в единый вычислительный комплекс для решения указанного типа задач. В этой связи необходимо решить проблему сопряжения модульных многопроцессорных систем для расширения вычислительных возможностей при решении сильносвязанных задач.

1.4. Анализ последних исследований и публикаций

Достаточно полный и всесторонний анализ оценок эффективности модульной многопроцессорной кластерной системы приведен в [8, 12 – 14].

Заметим, что в многопроцессорных системах в течение каждой итерации процессоры обмениваются данными на стыках вычислительных областей, используя актуальные значения переменных. В этой связи в приведенных работах рассматривались особенности пересылки данных между соседними узлами многопроцессорной вычислительной системы. При этом сравнивались варианты одностороннего и двустороннего режима передачи данных между процессорами. Как и ожидалось, при двустороннем режиме обмена данных уменьшилось ускорение вычислений за счет увеличения времени граничного обмена данных.

С целью исследования путей ускорения вычислений отдельно исследовалось влияние способа передачи данных между узлами кластерной

системы. В этой связи анализировалось влияние полудуплексного и дуплексного режима передачи данных в вычислительных сетях. Исследования показали, 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 отмеченные обстоятельства, в указанных работах подчеркивается важность исследования характеристик эффективности многопроцессорной системы при реализации дуплексного режима работы. Исходные данные для этого режима исследований представлены в табл. 1.1.

Таблица 1.1. Исходные данные для расчета характеристик эффективности многопроцессорной системы при реализации дуплексного режима работы кластера

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

Здесь приняты следующие обозначения: V_p – протокольная пропускная способность сети кластера, Гбит/с; T_{it} – время счета одной итерации относительно области вычислений, с; R – объем оперативной памяти узла кластера, Гбит; значение m может равняться единице для одностороннего режима граничного обмена данными, или двум для двустороннего; d – полудуплексный (d = 1) или дуплексный (d = 2) режим работы вычислительной сети кластерной системы.

Полученные результаты моделирования сведены в табл. 1.2.

Таблица 1.2. Результаты расчета основных характеристик эффективности многопроцессорной системы при реализации дуплексного режима работы кластера

Колич. узлов, N	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. 1.2, 1.3).

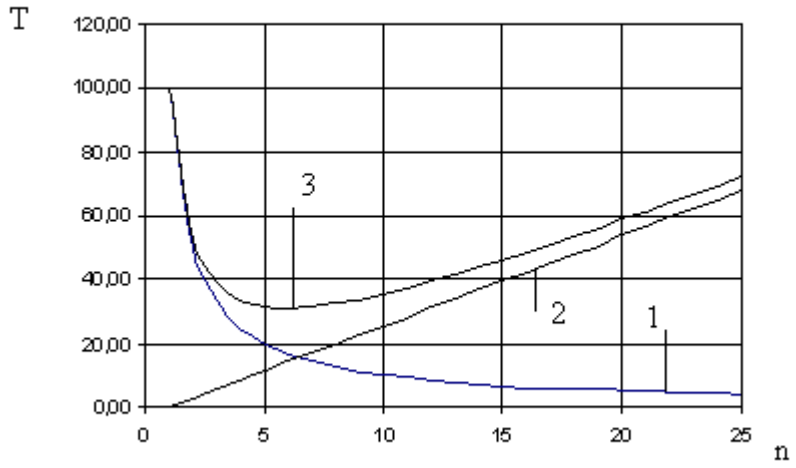

Рис. 1.2. Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы при дуплексном режиме работы

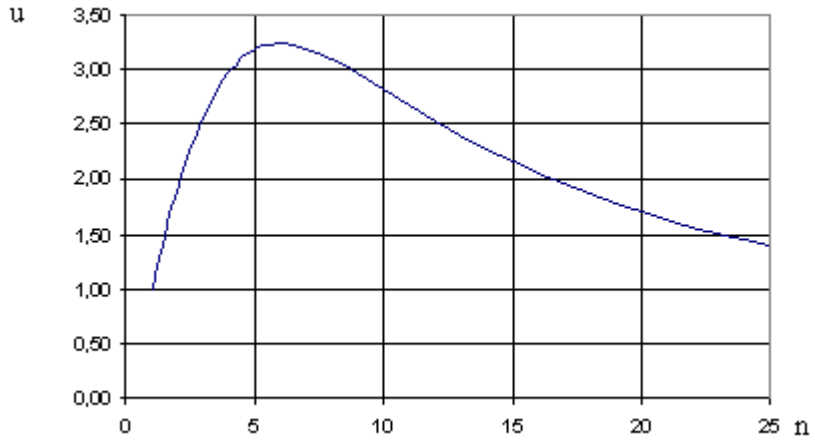

Рис. 1.3. Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы при дуплексном режиме работы

Проведенный анализ полученных результатов моделирования показал следующее. Как видно из рис. 1.2, время счета одной итерации при увеличении числа узлов многопроцессорной системы уменьшается по гиперболической зависимости (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 превышает шесть (рис. 1.2), то общее время счета начинает расти. Происходит это на фоне увеличения объема данных, которые пересылаются между узлами кластера.

С другой стороны, сравнительный анализ полудуплексного и дуплексного режимов работы системы показал, что в режиме “дуплекс” существенно уменьшилось время вычислений. Кроме того, значительно возросла характеристика ускорения вычислений [3].

Проведенный обзор показывает, что, с одной стороны, вопросам исследования эффективности многопроцессорных кластерных систем уделяется

достаточно серьезное внимание. С другой стороны, такие исследования охватывают самые разнообразные режимы функционирования кластерных многопроцессорных систем. Однако можно отметить, что на сегодняшний день проблема влияния сетевого интерфейса кластерной системы на оценки ее эффективности раскрыта не полностью. Более того, также не получило развития направление исследований, связанное с совершенствованием архитектуры сетевого интерфейса с целью повышения скорости обмена данных между вычислительными узлами многопроцессорной кластерной системы. Проведенные в данной работе исследования направлены на устранение указанного недостатка при анализе эффективности многопроцессорных кластерных систем и в дальнейшем развивают подход к исследованию таких систем, освещенный авторами в [3, 12 – 14].

1.5. Исследование основных режимов функционирования сетевого интерфейса кластерной системы

На первом этапе исследований рассмотрим особенности формирования архитектуры сетевого интерфейса кластерной системы и основные режимы его работы, а затем, уже на втором этапе, проведем анализ взаимодействия процессоров кластерной системы с его интерфейсом.

Итак, для оценки процессов, протекающих в кластерной системе при организации соответствующих потоков информации, необходимо сравнивать пропускную способность сети кластера и пропускную способность коммутатора. ***Эта процедура необходима для оптимального подбора компонентов сетевого интерфейса многопроцессорной кластерной системы***. В этой связи, для удобства исследований введем в рассмотрение такой параметр, как общая пропускная способность сети кластерной системы по спецификации производителя (V_s):

$$V_s = V_p \cdot N \cdot d. \qquad (1.1)$$

Здесь N – число узлов кластера, а V_p – протокольная пропускная способность сети кластера, Гбит/с.

Отметим, что современные коммутаторы имеют такую характеристику как “пропускная способность шины”. Например, двенадцатипортовый коммутатор 3*Com SuperStack* 3 *Switch* 3812 производителя 3*COM LG* имеет пропускную способность 24 Гбит/с. Это означает, что каждый из двенадцати портов может одновременно с другими в дуплексном режиме передавать и принимать данные с суммарной скоростью 2 Гбит/с. Точнее говоря, пропускная способность шины задана в пакетах в секунду, и заявленные 24 Гбит/с достигаются при пересылке больших пакетов. Будем считать, что рассматриваемые граничные области обладают именно такими характеристиками.

При таком подходе уже можно сравнивать общую пропускную способность сетевого интерфейса кластера (V_s) и пропускную способность коммутатора (V_b). Очевидно, что пропускная способность используемого коммутатора (V_b) будет соответствовать 24 Гбит/с. Для дальнейшего анализа сетевого интерфейса кластерной системы введем в рассмотрение коэффициент k_s и будем его трактовать как коэффициент пропускной способности сети кластера, который определим следующем образом:

$$k_s = \frac{V_s}{V_b}. \qquad (1.2)$$

С учетом формулы (1.1), получаем:

$$k_s = \frac{V_p \cdot N \cdot d}{V_b}. \qquad (1.3)$$

В работе [3] проведенные исследования справедливы для так называемого “идеального” кластера, когда k_s=1, т.е. вся необходимая информация для пересылок в кластерной системе через коммутатор распределяется без задержек между необходимыми ее узлами. Такой поход для исследования

процессов, протекающих в кластерных системах, оказался полезным для того, чтобы выявить основные режимы работы кластерной системы и соответственным образом их оценить. Однако на практике “идеальный” кластер создать практически невозможно. В этой связи, ранее выявленные особенности функционирования кластерной системы необходимо уточнить с учетом новых открывшихся обстоятельств. Для развития такого подхода введем в рассмотрение понятие коэффициента пропускной способности коммутатора (k_b):

$$k_b = \frac{V_b}{V_p \cdot N \cdot d}. \qquad (1.4)$$

Кроме того, для анализа согласования выбранной коммутационной шины с возможностями коммутатора введем в рассмотрение коэффициент полосы пропускания коммутатора (c_k), который будет определяться соотношением вида:

$$c_k = \frac{V_b}{N}. \qquad (1.5)$$

Для полной картины проводимых исследований введем некоторые определения, а затем, с учетом сформулированных определений, проведем более детальный анализ основных сетевых характеристик кластерной системы.

Определение 1.1. *Точкой сетевого равновесия* будем называть точку, в которой значения коэффициента пропускной способности сети кластерной системы и коэффициента пропускной способности коммутатора будут равны между собой.

Определение 1.2. *Равновесным числом узлов кластерной системы* будем называть то количество узлов, которое соответствует точке сетевого равновесия.

Определение 1.3. *Идеальной многопроцессорной кластерной системой* будем называть систему, в которой справедливо равенство вида $k_s = k_b$.

Определение 1.4. *Реальной многопроцессорной кластерной системой* будем называть систему, в которой справедливо неравенство вида $k_s \neq k_b$.

Определение 1.5. *Режимом дефицита сетевого интерфейса кластерной системы* будем называть вариант функционирования сети кластера, при котором выполняется неравенство вида $k_s < k_b$.

Определение 1.6. *Режимом профицита сетевого интерфейса кластерной системы* будем называть случай функционирования сети кластера, при котором выполняется неравенство вида $k_s > k_b$.

В рамках анализа работы сетевого интерфейса кластерной системы на первом этапе отметим некоторые необходимые особенности функционирования коммутатора. При этом заметим, что производительность коммутатора существенно зависит от типов коммутации. Применяемый в исследуемой многопроцессорной системе коммутатор поддерживает четыре типа коммутации:

– сквозная коммутация (*cut-through*);

– коммутация с буферизацией (*store-and-forward switching*);

– бесфрагментная коммутация (*fragment-free switching*);

– адаптивная коммутация (*intelligent*).

При сквозной коммутации в буфер входного порта поступают лишь несколько первых байтов пакета, что необходимо для считывания адреса назначения. После установления адреса назначения, параллельно с приемом остальных байтов кадра, происходит коммутация необходимого маршрута, и пакет передается к выходному порту, если он не используется другими устройствами кластера. В противном случае, весь пакет поступает в буфер входного порта. Сквозная коммутация обеспечивает самую высокую скорость коммутации, что дает значительный выигрыш в производительности.

При коммутации с буферизацией пакет поступает в буфер входного порта, где по контрольной сумме проверяется на наличие ошибок. Если ошибки не обнаружены, пакет передается на выходной порт. Этот способ коммутации гарантирует фильтрацию от ошибочных пакетов, однако за счет процедуры

буферизации снижается пропускная способность коммутатора по сравнению со сквозной коммутацией.

При бесфрагментной коммутации в буфер входного порта поступает не весь пакет, а только первые 64 байта. Для пакета минимального размера это соответствует полной буферизации, а для пакетов, размер которых больше 64 байт, это соответствует сквозной коммутации. Таким образом, при бесфрагментной буферизации проверке подлежат только кадры минимального размера.

При адаптивной коммутации коммутатор сам выбирает для каждого порта оптимальный режим работы.

Предварительно заметим, что для получения высоких оценок эффективности и ускорения вычислений кластерной системы необходимо, чтобы она функционировала в одном из указанных типов коммутации в режиме дефицита сетевого интерфейса. Такое утверждение вполне очевидно, т.к. суммарная скорость данных, передаваемых и принимаемых всеми узлами многопроцессорной кластерной системы, не должна превышать пропускной способности коммутатора.

Основная особенность режима профицита сетевого интерфейса состоит в том, что коммутатор может сталкиваться с перегрузками, когда сумма входящих трафиков превышает пропускную способность коммутационной матрицы коммутатора. В таком случае меняются условия коммутации данных. Здесь коммутатор переходит в режим коммутации с буферизацией, что приводит к уменьшению производительности коммутатора. Однако критическое увеличение объема данных в буфере порта приводит к его переполнению и, как следствие, к простою коммутатора. При указанных обстоятельствах коммутатор не сможет обеспечивать максимально стабильное и надежное формирование коммутируемых потоков данных в вычислительной системе. Основные характеристики такой кластерной системы (в т.ч. и быстродействие) будут существенно ухудшаться, а основная задача формирования многопроцессорности вычислений теряет свой смысл.

При отмеченных обстоятельствах выполним процедуру моделирования указанных коэффициентов в зависимости от количества узлов кластерной системы.

Исходные данные для изучения области изменения коэффициентов сетевого интерфейса многопроцессорной системы перечислены в табл. 1.3.

Таблица 1.3. Исходные данные для расчета сетевых характеристик кластерной системы

V_p	1 Гбит/с
V_b	24 Гбит/с

На первом этапе исследований выведем аналитическое соотношение для определения равновесного числа узлов кластерной системы. Для этой цели приравняем коэффициенты $k_s = k_b$:

$$\frac{V_p \cdot N \cdot d}{V_b} = \frac{V_b}{V_p \cdot N \cdot d}. \qquad (1.6)$$

После некоторых преобразований соотношения (1.6) получаем квадратное уравнение, требуемое значение корня которого будет определяться соотношением вида:

$$N = \frac{V_b}{V_p \cdot d}. \qquad (1.7)$$

Анализ соотношения (1.7) показывает, что равновесное число узлов кластерной системы зависит не только от протокольной пропускной способности сети кластера и пропускной способности коммутатора, но и от режима работы вычислительной сети кластерной системы (полудуплекс, дуплекс). Дуплексный режим работы вычислительной сети в два раза уменьшает равновесное число узлов кластерной системы. Такой результат вполне очевиден, т.к. благодаря дуплексному режиму работы сетевого

интерфейса в два раза расширяется общая пропускная способность сети кластерной системы.

Итак, для рассматриваемой кластерной системы, с учетом заявленных возможностей сетевого интерфейса (табл. 1.3), на основании соотношения (1.7) определим равновесное число узлов кластерной системы. При этом равновесное число узлов кластерной системы соответствует $N = 12$. Выполним анализ полученного значения. Очевидно, что в рамках рассматриваемых параметров сетевого интерфейса “идеальной” кластерная система будет в случае, когда число ее узлов соответствует $N = 12$. Заметим, что это полностью согласуется с техническими возможностями выбранного коммутатора (количество его портов равно двенадцати). На этом этапе исследований можно отметить, что для рассматриваемого режима работы вычислительной сети коммутатор подобран удачно.

Кроме того, приведенные результаты расчета равновесного числа узлов кластерной системы убедительно опровергают утверждение о том, что "чем больше узлов в кластерной системе, тем быстрее она работает". Так, при $N > 12$ кластерная система переходит в режим профицита сетевого интерфейса, который характерен для малопроизводительных многопроцессорных систем, сетевой интерфейс которых может быть построен на примитивных коммутаторах или вообще при его неудачной организации.

Далее, для уточнения особенностей функционирования сетевого интерфейса кластерной системы, на основании соотношений (1.3) – (1.5) выполним процедуру моделирования основных его числовых характеристик.

Полученные результаты моделирования сведены в табл. 1.4.

Таблица 1.4. Результаты расчета основных сетевых коэффициентов кластерной системы

Колич. узлов, N	k_s	k_b	c_k
1,00	0,08	12,00	24,00
2,00	0,17	6,00	12,00
3,00	0,25	4,00	8,00
4,00	0,33	3,00	6,00
5,00	0,42	2,40	4,80
6,00	0,50	2,00	4,00
7,00	0,58	1,71	3,43
8,00	0,67	1,50	3,00
9,00	0,75	1,33	2,67
10,00	0,83	1,20	2,40
11,00	0,92	1,09	2,18
12,00	1,00	1,00	2,00
13,00	1,08	0,92	1,85
14,00	1,17	0,86	1,71
15,00	1,25	0,80	1,60
16,00	1,33	0,75	1,50
17,00	1,42	0,71	1,41
18,00	1,50	0,67	1,33
19,00	1,58	0,63	1,26
20,00	1,67	0,60	1,20
21,00	1,75	0,57	1,14
22,00	1,83	0,55	1,09
23,00	1,92	0,52	1,04
24,00	2,00	0,50	1,00
25,00	2,08	0,48	0,96

Результаты моделирования представлены также в виде графических зависимостей (рис. 1.4). Для удобства анализа исследуемых коэффициентов соответствующие графические зависимости представлены в одной системе координат.

Проведем предварительный анализ полученных результатов. Очевидно, что с увеличением числа узлов кластерной системы будет возрастать пропускная способность сети кластера (V_s). Тогда изменение коэффициента пропускной способности сети кластера (k_s, формула 1.3) будет осуществляться по линейному закону (рис. 1.4, линия 1). С другой стороны, увеличение объема данных, пересылаемых между узлами кластера, приведет к тому, что коммутатор будет перегружаться, и его коэффициент пропускной способности

(k_b, формула 1.4) будет уменьшаться по нелинейному закону (рис. 1.4, линия 2).

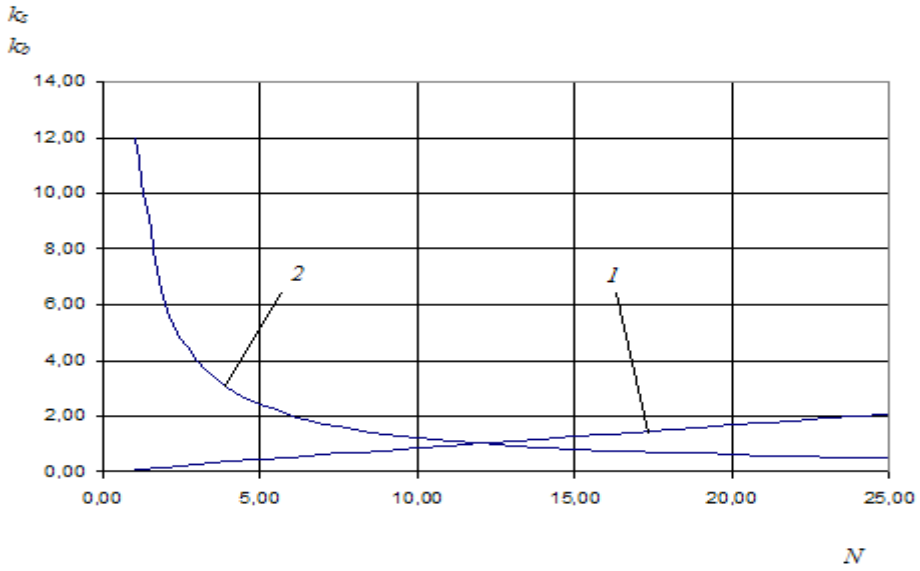

Рис. 1.4. Зависимости основных сетевых коэффициентов кластерной системы от количества узлов

Далее перейдем к более детальному анализу результатов моделирования основных сетевых характеристик кластерной системы с учетом отмеченных обстоятельств. При этом вновь рассмотрим некоторые особенности работы коммутатора. Так, если N узлов многопроцессорной системы пытаются установить соединение с одним узлом по технологии *GigabitEthernet*, то коммутационная шина коммутатора может выделить каждому узлу лишь полосу пропускания (c_k), которая будет определяться соотношением (1.5). В табл. 1.5 приведен расчет полосы пропускания коммутатора (c_k) для такого режима работы сетевого интерфейса. Заметим, что для N = 12 значение полосы пропускания на каждый выходной порт кластерной системы будет соответствовать 2 Гбит/с, а это полностью согласуется как с дуплексным режимом обмена данных в многопроцессорной системе, так и с возможностями самой коммутационной шины. При таких обстоятельствах сетевой интерфейс многопроцессорной системы будет функционировать в режиме его дефицита и в условиях максимально допустимой загрузки каналов коммутации коммутатора. Таким образом, технические возможности заявленного

коммутатора будут полностью согласовываться с возможностями коммутационной сети.

Итак, возникают предпосылки для переоценки характеристик ускорения и эффективности кластерной системы. На первом этапе введем такую характеристику, как коэффициент пропускной способности кластерной системы (k_k):

$$k_k = \begin{cases} 1, \text{для режима дефицита сетевого интерфейса и} \\ \quad \text{для идеального кластера,} \\ k_b, \text{для режима профицита сетевого интерфейса.} \end{cases} \qquad (1.8)$$

Зависимость коэффициента пропускной способности кластерной системы от числа узлов представлена в графическом виде на рис. 1.5. Выполним его анализ.

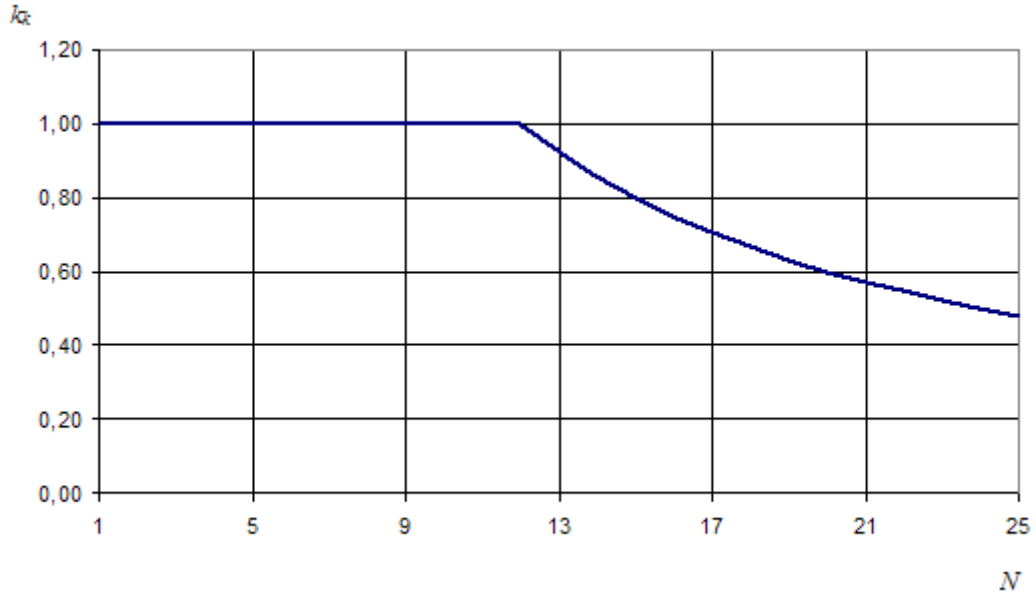

Рис. 1.5. График зависимости коэффициента пропускной способности кластерной системы от числа узлов

Вполне очевидно, что для режима дефицита сетевого интерфейса коэффициент пропускной способности кластерной системы будет определяться характеристиками сетевого интерфейса. При этом в силу особенностей режима дефицита сетевого интерфейса, такой коэффициент принимают равным единице, т.к. коммутационная матрица будет работать в режиме сквозной

коммутации (*cut-through*) и информация будет передаваться без использования процедуры буферизации передаваемых пакетов с наибольшей скоростью. В режиме профицита сетевого интерфейса такой коэффициент будет определяться характеристиками коммутатора, когда сумма входящих трафиков превышает пропускную способность коммутационной матрицы коммутатора. Здесь коммутатор переходит в режим коммутации с буферизацией, что приводит к потере его производительности, это обстоятельство и отражено убывающей линией на рис. 1.5.

1.6. Влияние сетевого интерфейса на эффективность модульной многопроцессорной кластерной системы

1.6.1. Особенности сопряжения вычислительных узлов с сетевым интерфейсом многопроцессорной системы

Для исследования влияния сетевого интерфейса кластерной системы на оценки эффективности рассмотрим особенности сопряжения вычислительных узлов кластерной системы с сетевым интерфейсом. На рис. 1.6 приведена схема сопряжения вычислительных узлов многопроцессорной системы с сетевым интерфейсом.

Особенности конфигурации сети состоят в следующем. Коммутатор *SW1* образует сеть управления, загрузки и диагностики расширенного кластера, все интегрированные интерфейсы мастер – узла и *slave* – узлов соединяются со входами/выходами этого коммутатора. Коммутатор *SW*1 образует сеть управления загрузки и диагностики кластера. Так, интегрированный сетевой интерфейс *PM*001.*i*01 узла *MNode*001 с функцией сетевой загрузки подсоединен входом/выходом к порту 11 управляемого коммутатора *SW*1. Мастер узел *MNode*001 соединен входом/выходом при помощи двунаправленных внешних сетевых интерфейсов *Gi*_001,1 и *Gi*_001,2 с сетью *Ethernet*.

Интегрированный сетевой интерфейс *PN*001.*i*01 узла *NNode*001 подсоединен входом/выходом к порту 01 управляемого коммутатора *SW*1.

Аналогично выполнено соединение остальных узлов кластерной системы с управляемым коммутатором *SW*1. В то же время, порт 12 управляемого коммутатора *SW*1 соединен с портом 12 управляемого коммутатора *SW*2 и используется для *Web* конфигурирования и диагностики коммутаторов.

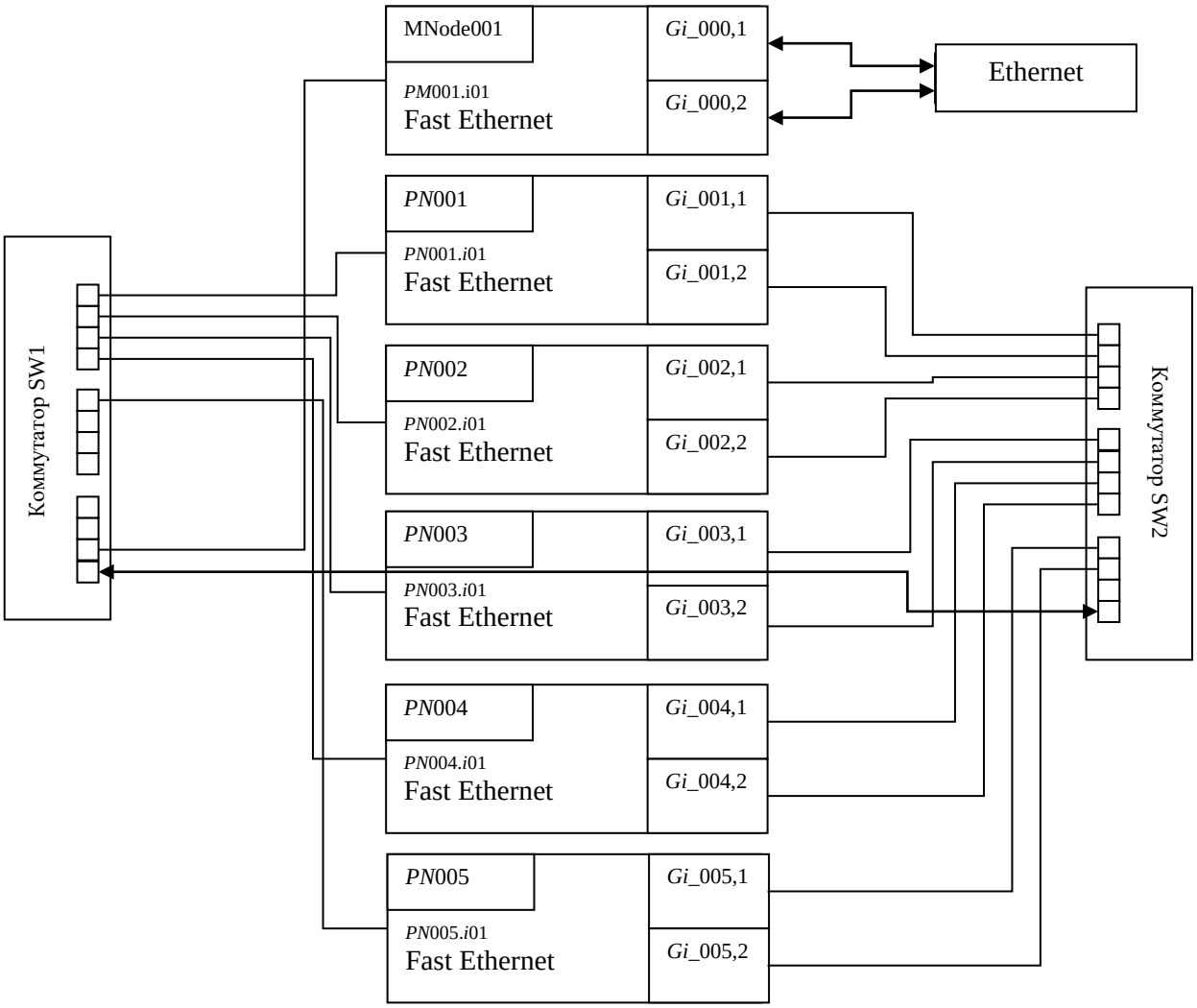

Рис. 1.6. Схема сопряжения вычислительных узлов многопроцессорной системы с сетевым интерфейсом GigabitEthernet

Архитектура вычислительной сети кластера реализована следующим образом. *Slave* – узел *PNode*001 соединен входом/выходом двунаправленным внешним сетевым интерфейсом *Gi*_001,1 с 01 портом управляемого коммутатора *SW*2. Кроме того, такой узел дополнительно соединен двунаправленным сетевым интерфейсом *Gi*_001,2 с 02 портом управляемого

коммутатора *SW*2. *Slave* – узел *PN*ode002 соединен входом/выходом двунаправленным внешним сетевым интерфейсом *Gi*_002,1 с 03 портом управляемого коммутатора *SW*2. Кроме того, такой узел дополнительно соединен двунаправленным сетевым интерфейсом *Gi*_002,2 с 04 портом управляемого коммутатора *SW*2. *Slave* – узел *PN*ode003 соединен входом/выходом двунаправленным внешним сетевым интерфейсом *Gi*_003,1 с 05 портом управляемого коммутатора *SW*2. Кроме того, такой узел дополнительно соединен двунаправленным сетевым интерфейсом *Gi*_003,2 с 06 портом управляемого коммутатора *SW*2. *Slave* – узел *PNode*004 соединен входом/выходом двунаправленным внешним сетевым интерфейсом *Gi*_004,1 с 07 портом управляемого коммутатора *SW*2. Кроме того, такой узел дополнительно соединен двунаправленным сетевым интерфейсом *Gi*_004,2 с 08 портом управляемого коммутатора *SW*2. *Slave* – узел *PNode*005 соединен входом/выходом двунаправленным внешним сетевым интерфейсом *Gi*_005,1 с 09 портом управляемого коммутатора *SW*2. Кроме того, такой узел дополнительно соединен двунаправленным сетевым интерфейсом *Gi*_005,2 с 10 портом управляемого коммутатора *SW*2.

Для решения широкого круга прикладных задач граничный обмен данных между вычислительными узлами целесообразно реализовать между соседними узлами. В таком случае связь между узлами кластера организуется по топологии кольцо (рис. 1.7), т.е. узел *PN*001 обменивается данными с *PN*002, узел *PN*002 с *PN*003, узел *PN*003 с *PN*004, узел *PN*004 с *PN*005, узел *PN*005 с *PN*001.

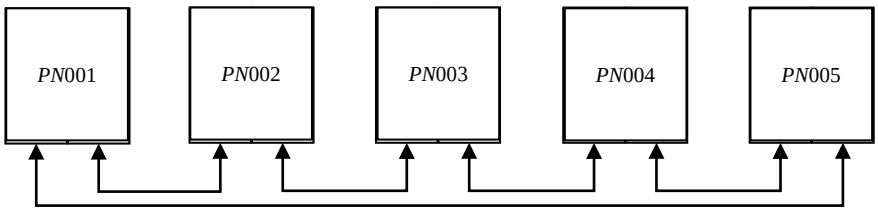

Рис. 1.7. Структура вычислительной сети кластера для реализации граничного обмена

Для исключения взаимного влияния при граничном обмене (передаче/приеме) данными между вычислительными узлами, создаются виртуальные локальные сети (*VLAN*), ограничивающие трафик в пределах отдельной *VLAN* сети коммутатора *SW*2. В таком коммутаторе сформированы пять виртуальных сетей: *vs*1.2 между портами 02 и 03 коммутатора *SW*2, *vs*2.3 между портами 04 и 05 коммутатора *SW*2, *vs*3.4 между портами 06 и 07 коммутатора *SW*2, *vs*4.5 между портами 08 и 09 коммутатора *SW*2, *vs*5.1 между портами 10 и 01 коммутатора *SW*2. Реализован режим работы системы с наличием промежуточного буфера памяти для хранения соответствующих пакетов. Промежуточный буфер памяти коммутаторов *SW*1 – *SW*2 избавляет от необходимости процедуры синхронизации данных при сетевом обмене, когда реализуется процесс отправки и приема пакетов для сильносвязанных задач. При этом возникает возможность уменьшить загрузку *CPU*, что повышает эффективность и производительность кластерной системы в целом.

1.6.2. Особенности подбора элементов сетевого интерфейса кластерной многопроцессорной системы

На первом этапе исследований рассмотрим особенности выбора архитектуры сетевого интерфейса кластерной системы и основные режимы его работы. Для удобства освещения такого подхода предварительно отметим, что вычислительная сеть кластерной системы имеет две основные характеристики – пропускную способность и латентность. Пропускная способность вычислительной сети – это скорость передачи данных между двумя ее узлами после того, как связь между ними установлена. Латентность – это среднее время между вызовом функции передачи данных и самой передачей. Такое время затрачивается на адресацию информации, срабатывание промежуточных сетевых устройств, а также других сетевых особенностей, возникающих при передаче данных.

Вообще заметим, что пропускная способность и латентность не только характеризуют кластер, но и ограничивают класс выполняемых задач. Так, если

задача требует интенсивного обмена данных пакетов небольшой длины, кластер, использующий сетевое оборудование с большой латентностью, будет большую часть времени тратить на установление сетевых соединений, а не на передачу данных между узлами системы. Следовательно, узлы в кластерной системе будут простаивать, и тогда эффективность распараллеливания будет существенно уменьшаться.

С другой стороны, если пересылаются большие пакеты данных, то влияние периода латентности на эффективность кластера может снижаться за счет того, что передача занимает гораздо больше времени, чем установление самого соединения.

Принимая во внимание отмеченное, рассмотрим особенности выбора сетевого интерфейса при конструировании модульной многопроцессорной кластерной системы [1, 2].

Сетевые кабели. Для сетевых соединений в многопроцессорной вычислительной системе применялась технология *GigabitEthernet* [15]. При этом технология 1000*BASE-T, IEEE* 802.3*ab* – стандарт, использующий витую пару категорий 5*e*. Такое сетевое оборудование можно рассматривать как недорогую альтернативу оптическим кабелям.

Сетевой адаптер. В качестве сетевых адаптеров можно использовать любые карты, поддерживающие работу в стандарте *GigabitEthernet*. В настоящее время существует много производителей, выпускающих такие сетевые адаптеры. В качестве примера можно привести следующие: *Compex, Intel, Macronix* и др. При проектировании рассматриваемой многопроцессорной системы предпочтение было отдано производителю 3*COM* [16].

Сетевые карты фирмы 3 *COM* [17], при равных прочих условиях, имеют некоторые преимущества, заметно влияющие на производительность сетевых коммуникаций. Укажем некоторые основные из них.

Разгрузка процессора при вычислении контрольных сумм сетевых пакетов данных. Такой режим работы сетевого адаптера освобождает

центральный процессор от интенсивных вычислений контрольных сумм, выполняя их в самой сетевой плате. Тем самым повышается производительность системы.

Освобождение ЦП при восстановлении сегментированных пакетов. Такое свойство сетевого адаптера позволяет снизить нагрузку на центральный процессор, повышая производительность системы.

Объединение прерываний позволяет группировать несколько полученных сетевых пакетов, сокращает число прерываний и максимально освобождает процессорные ресурсы для работы приложений.

Принимая во внимание отмеченное, для рассматриваемой кластерной системы была выбрана сетевая карта производителя 3*COM* 996*B-T Server adapter* [17].

Коммутаторы. Одним из важных элементов сетевого интерфейса кластерной системы являются устройства коммутации сетевых каналов. Для рассматриваемой многопроцессорной системы был выбран двенадцатипортовый коммутатор 3*Com SuperStack* 3 *Switch* 3812 производителя 3*COM LG* [18], который имеет пропускную способность 24 Гбит/с. Это означает, что каждый из двенадцати портов может одновременно с другими в дуплексном режиме передавать и принимать данные с суммарной скоростью 2 Гбит/с. Такой коммутатор относится к семейству коммутаторов, предназначенных для построения гигабитных сетей на основе медных соединений. Они поддерживают стандартный набор сетевых технологий: виртуальные сети, приоритезация трафика, агрегированные каналы, фильтрация многоадресного трафика и другие.

Технические характеристики сетевого оборудования освещаемой многопроцессорной кластерной системы приведены в табл. 1.5.

Таблица 1.5. Технические характеристики сетевого оборудования кластерной системы

Сетевой кабель	Тип	*GigabitEthernet*
	Пропускная способность	1 Гбит/с
	Стандарт	5*e*
	Цена	$ 6
Сетевой адаптер	Тип	996*B-T Server adapter*
	Производитель	3*COM*
	Пропускная способность	1 Гбит/с
	Цена	$ 90
Коммутатор	Тип	3*Com SuperStack* 3 *Switch* 3800
	Производитель	3*COM*
	Пропускная способность	24 Гб/с
	Цена	$ 940

При указанных сетевых характеристиках кластерной системы были проведены соответствующие вычислительные эксперименты, а также исследованы характеристики эффективности многопроцессорной системы.

1.6.3. Исследование оценок эффективности кластерной системы

Исходные данные для исследования оценок эффективности кластерной системы представлены в табл. 1.6.

Таблица 1.6. Исходные данные для расчета оценок эффективности многопроцессорной системы

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

Моделирование основных оценок эффективности кластерной системы выполнено в соответствии с аналитическими соотношениями, выведенными в работе [14].

Полученные результаты моделирования сведены в табл. 1.7.

Таблица 1.7. Результаты расчета основных оценок эффективности многопроцессорной системы

Колич. узлов, N	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 основных характеристик эффективности многопроцессорной системы представлены также в виде графических зависимостей (рис. 1.8 и 1.9).

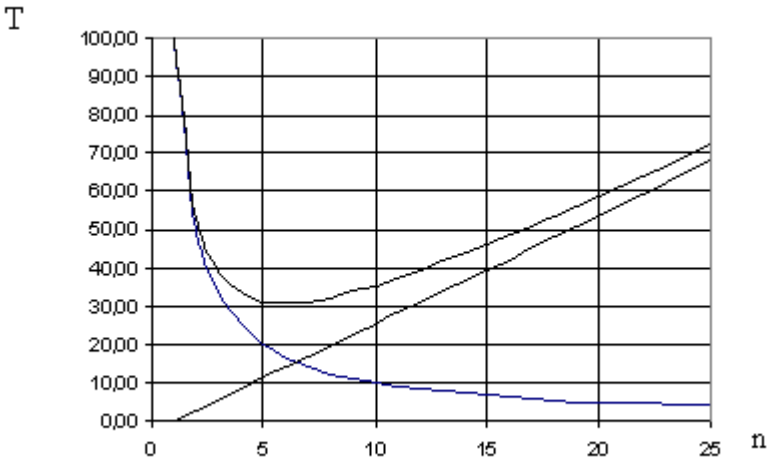

Рис. 1.8. Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы при дуплексном режиме работы

Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы представлена на рис. 1.9.

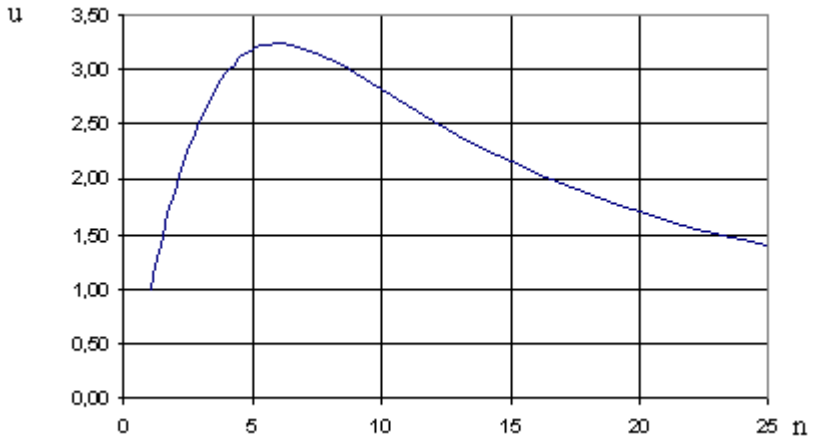

Рис. 1.9. Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы

Уточним оценки эффективности кластерной системы с учетом влияния сетевого интерфейса. Тогда, принимая во внимание, что скорость порта узла кластерной системы равна протокольной пропускной скорости сети, общая оценка пропускной способности сети кластера будет определяться следующим соотношением:

$$V = d \cdot V_p \cdot k_k . \qquad (1.9)$$

Очевидно, что с учетом выражения (1.6) такая формула будет распадаться на две: одна описывает общую оценку пропускной способности сети кластера для режима дефицита сетевого интерфейса (V_1), а другая – режима его профицита (V_2). Для режима дефицита сетевого интерфейса получим:

$$V_1 = d \cdot V_p , \qquad (1.10)$$

для режима профицита сетевого интерфейса такая скорость будет выражаться формулой вида:

$$V_2 = \frac{V_b \cdot d}{N} . \qquad (1.11)$$

Анализ соотношения (1.10) показывает, что для режима дефицита сетевого интерфейса общая оценка пропускной способности сети кластера будет зависеть от скорости порта узла многопроцессорной сети и режима передачи данных в вычислительной сети (дуплекс или полудуплекс). В то же

время такая скорость не будет зависеть от числа узлов многопроцессорной системы. Этот, на первый взгляд, парадоксальный вывод можно объяснить тем, что система работает в режиме дефицита сетевого интерфейса, а это означает, что до равновесного числа узлов кластерной системы скорость коммутации данных в сети будет определяться, в основном, скоростью узла порта многопроцессорной системы.

С другой стороны, в режиме профицита сетевого интерфейса оценка пропускной способности сети кластера будет зависеть от пропускной способности используемого коммутатора, режима передачи данных в вычислительной сети (дуплекс или полудуплекс) и количества вычислительных узлов кластерной системы.

В таком случае соотношение для определения времени граничного обмена данных между узлами кластерной системы будет распадаться на два:

$$T_{ex1} = \frac{E}{V_1} \tag{1.12}$$

для работы кластерной системы в режиме дефицита сетевого интерфейса и

$$T_{ex2} = \frac{E}{V_2} \tag{1.13}$$

для работы кластерной системы в режиме профицита сетевого интерфейса.

В выражениях (1.12, 1.13) E – объем данных в области граничного обмена (Гбит). Выразим соотношения (1.12) и (1.13) через параметры кластерной системы:

$$T_{ex1} = \frac{m \cdot (N-1) \cdot \sqrt{R}}{d \cdot V_p}, \tag{1.14}$$

$$T_{ex2} = \frac{m \cdot (N-1) \cdot \sqrt{R} \cdot N}{d \cdot V_b}. \tag{1.15}$$

Далее, для изучаемой кластерной многопроцессорной системы в условиях рассматриваемого эксперимента оценим количество узлов кластерной системы, при котором задача будет решаться наиболее эффективно. При этом заметим,

что время счета одной итерации вычислительного процесса складывается из двух слагаемых – времени непосредственного счета на процессорах $T_{calc} = \frac{T_{it}}{N}$ и времени обмена данных между вычислительными узлами кластера T_{ex}, т.е.

$$T_{it} = T_{calc} + T_{ex}. \tag{1.16}$$

При этом в [3, 19] показано, что скорость вычислений будет расти примерно до момента, когда

$$T_{calc} \approx T_{ex}. \tag{1.17}$$

Таким образом, исходя из соотношения (1.17), предоставляется возможность оценить количество узлов кластерной вычислительной системы, при котором задача будет решаться наиболее эффективно. Заметим, что данный этап предусматривает ту цель исследований, при которой процесс распараллеливания направлен на уменьшение общего времени счета. Очевидно, что при этом общий размер разностной сетки не зависит от числа вычислительных узлов кластерной системы. Учитывая соотношение (1.17), получают аналитические выражения для определения оптимального числа узлов кластерной системы:

$$\frac{T_{it}}{N} \approx \frac{m \cdot (N-1) \cdot \sqrt{R}}{d \cdot V_p} \tag{1.18}$$

для работы кластера в режиме дефицита сетевого интерфейса и

$$\frac{T_{it}}{N} \approx \frac{m \cdot (N-1) \cdot \sqrt{R} \cdot N}{d \cdot V_b} \tag{1.19}$$

для работы кластера в режиме профицита сетевого интерфейса. На основании выражений (1.18) и (1.19) можно получить два уравнения относительно N для определения оптимального числа узлов кластерной системы, при котором общее время вычислений, требуемое для решения задачи, будет минимальным.

Уравнение (1.18) преобразуется к квадратичному виду

$$N^2 - N - \frac{T_{it} \cdot d \cdot V_p}{m \cdot \sqrt{R}} = 0 . \qquad (1.20)$$

Решением такого уравнения будет два корня, при этом один из них положительный, а другой – отрицательный. Исходя из поставленных физических условий задачи, принимается положительный корень, значение которого равно шести, т.е. N = 6. Заметим, что такое решение удовлетворяет неравенству из определения 1.5, которое устанавливает условия функционирования кластерной системы в режиме дефицита сетевого интерфейса.

Уравнение (1.19) будет сводиться к кубическому виду

$$N^3 - N^2 - \frac{T_{it} \cdot d \cdot V_b}{m \cdot \sqrt{R}} = 0 \qquad (1.21)$$

и оно будет иметь два мнимых корня и один действительный. Действительный корень соответствует N = 9. Однако анализ такого корня показывает, что он не удовлетворяет условию функционирования кластерной системы в режиме профицита сетевого интерфейса (определение 1.6).

Проведенный анализ полученных результатов моделирования показал следующее. Наибольшая величина ускорения вычислений достигается на шести узлах многопроцессорной системы, а ориентировочная цена сетевого оборудования будет составлять около 3000 у.е.

Время счета задачи уменьшается со 100 с до 30,81 с. При выбранном размере кластера задача будет решаться в 3,25 раза быстрее, чем на одном компьютере.

1.6.4. Исследование загрузки вычислительной сети кластерной системы

Итак, рассмотрим характеристику коэффициента использования сети многопроцессорной кластерной системы. *Такая характеристика необходима для проверки правильности подобранного сетевого оборудования.* С этой

целью выведем соотношение для коэффициента использования сети через параметры кластерной системы. При этом принимается:

$$\xi = \frac{T_{ex1}}{T}. \qquad (1.22)$$

Такая характеристика будет рассчитываться для режима дефицита сетевого интерфейса, т.к. кластерная система ориентирована на работу в этом режиме. Выведем коэффициент использования сети кластерной системы через ее параметры.

Учитывая выражения (1.10), (1.16), окончательное значение коэффициента использования сети можно записать в виде аналитического соотношения, выраженного через параметры вычислительного кластера, т.е.:

$$\xi = \frac{m \cdot N \cdot (N-1) \cdot \sqrt{R}}{T_i \cdot d \cdot V_p + N \cdot m \cdot (N-1) \cdot \sqrt{R}}. \qquad (1.23)$$

Результаты расчета коэффициента использования сети для дуплексного режима работы многопроцессорной системы приведены в табл. 1.8.

Таблица 1.8. Результаты расчета коэффициента использования сети кластера

Колич. узлов	*КЗС*
1	0,00
2	0,05
3	0,15
4	0,25
5	0,36
6	0,46
7	0,54
8	0,61
9	0,67
10	0,72
11	0,76
12	0,79
13	0,82
14	0,84
15	0,86

Полученные результаты позволяют сделать вывод, что, как и ожидалось, при увеличении числа узлов кластерной системы значения коэффициента использования сети будут расти. С другой стороны, известно [20], что для сетевой технологии *Ethernet,* когда ξ = 50 % 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. При этом сеть многопроцессорной системы будет работать в режиме сквозной коммутации.

Таким образом, при загрузке сети до 50 % технология *Ethernet* на разделяемом сегменте хорошо справляется с передачей трафика, генерируемого узлами многопроцессорной системы. Однако при повышении интенсивности генерируемого трафика сеть все больше времени начинает проводить неэффективно, повторно передавая кадры, которые вызвали коллизию. При возрастании интенсивности генерируемого трафика возникает ситуация, когда практически любой кадр, который пытается передать некоторый узел многопроцессорной системы, сталкивается с другими кадрами, вызывая коллизию. Сеть перестает передавать полезную информацию и работает “на себя”, обрабатывая коллизии.

Следовательно, можно отметить, что при выбранном режиме функционирования кластера можно использовать не больше шести лезвий и сделать вывод, что оборудование сетевого интерфейса кластерной системы подобрано удачно.

1.7. Многоканальный режим функционирования сетевого интерфейса многопроцессорной системы

Анализ выявленных режимов функционирования сетевого интерфейса многопроцессорной системы позволяет сформулировать следующую проблему: каким образом можно расширить область дефицита сетевого интерфейса многопроцессорных систем или, иными словами, *как за счет конструктивных особенностей архитектуры вычислительных сетей многопроцессорных*

кластерных систем можно повысить оценки ее эффективности и быстродействия?

Такая проблема может быть решена следующим образом. Для увеличения пропускной способности сети кластера рекомендуется применять процедуру "связывания каналов" или технологию *channel bonding*. Технология связывания каналов (*channel bonding*) позволяет объединять несколько сетевых адаптеров в один скоростной канал. Предлагаемая сетевая архитектура многопроцессорной кластерной системы должна позволять, во-первых, повысить быстродействие вычислений во время решения сильносвязанных задач и, во-вторых, обеспечить высокоскоростной доступ к памяти узлов кластера, снижая загрузку канала, который проходит между узлами вычислительной системы.

1.7.1. Особенности сопряжения вычислительных узлов с сетевым интерфейсом для режима агрегации каналов сетевого интерфейса многопроцессорной системы

Применяемая технология связывания каналов сетевого интерфейса многопроцессорной кластерной системы позволяет объединить узлы кластера в сеть таким образом, чтобы каждый узел многопроцессорной системы подсоединялся к коммутатору более чем одним каналом. Для реализации такого подхода необходимо оснастить узлы кластера либо несколькими сетевыми платами, либо многопортовыми платами. Связывание каналов аналогично режиму транкинга при соединении коммутаторов, который используется для увеличения скорости передачи данных между двумя или несколькими коммутаторами. Применение процедуры связывания каналов под управлением ОС *Linux* позволяет организовать равномерное распределение нагрузки (приема/передачи данных) между соответствующими каналами многопроцессорной системы и увеличить скорость обмена данных между ее узлами. При этом необходимо отметить, что основная особенность такого режима работы кластерной системы состоит в том, что повышается надежность

ее функционирования. Так, в случае отказа адаптера трафик посылается через другие работающие адаптеры без прерывания обмена данных. Если же адаптер вновь начинает работать, то через него опять пересылаются данные.

Вообще заметим, что технология *агрегации каналов* может порождать некоторые проблемы, связанные с выбором коммутаторов и их настройкой. Приведем основные. Так, коммутатор должен поддерживать режим связывания каналов, иначе могут иметь место всевозможные ошибки при построении коммутатором таблиц маршрутизации пакетов или таблиц *MAC*-адресов. Такие коммутаторы должны поддерживать для своих портов функции *Link Aggregation* или соответствовать стандарту *IEEE 802.3ad*.

Другим вариантом реализации технологии агрегации каналов сетевого интерфейса может быть выбор коммутатора с возможностью поддержки режима виртуальных локальных сетей (*VLAN*). Применение *VLAN* призвано помочь избежать "дублирования" во внутренних таблицах коммутаторов *MAC*-адресов многопортовых сетевых плат. Впрочем, практика показывает, что и поддержка режима *VLAN* не всегда помогает эффективно разделить каналы.

Вообще заметим, что возможен и иной прием формирования многоканального сетевого интерфейса. Так, в [21] показано, что можно отказаться от применения специализированного сетевого оборудования, поддерживающего связывание каналов. Сетевые каналы при этом можно организовать при помощи двойного (тройного и т.д.) набора обычных хабов или свитчей. Здесь непересекающиеся сетевые сегменты организуются таким образом, чтобы каждый новый сетевой канал образовывал свою собственную сеть, физически не связанную с сетями других каналов.

Несмотря на широкий выбор методов формирования многоканального режима функционирования сетевого интерфейса в многопроцессорных вычислительных системах, приведем некоторые принципиальные особенности, которые следует иметь в виду при конструировании режима агрегации сетевого интерфейса. Так, все процессоры в подсети должны быть объединены одинаковым способом. Объединение каналов требует, как минимум, двух

физических подсетей, которых, тем не менее, может быть и больше. В версиях ядра ОС `2.4.x` технология *Channel Bonding* является стандартной опцией. Сетевые карты настраиваются, как обычно, за исключением того, что команду *ifconfig* необходимо применять для настройки сетевых карт в связке. Утилита *ifenslave* используется для активации оставшихся сетевых карт связанного соединения. Сети с объединенными каналами могут соединяться с обычными сетями посредством маршрутизатора или моста, поддерживающего технологию *Channel Bonding*.

Схема организации сетевого интерфейса по технологии *channel bonding* для рассматриваемой многопроцессорной системы приведена на рис. 1.10.

На первом этапе освещения сетевого интерфейса рассмотрим особенности конфигурации сети 1. Как и в предыдущей конструкции многопроцессорной кластерной системы коммутатор *SW*1 образует сеть управления, загрузки и диагностики кластера. Так, интегрированный сетевой интерфейс *PM*001.*i*01 узла *MNode*001 с функцией сетевой загрузки подсоединен входом/выходом к порту 01 управляемого коммутатора *SW*1. Интегрированный сетевой интерфейс *PN*001.*i*01 узла *PNode*001 с функцией сетевой загрузки подсоединен входом/выходом к порту 02 управляемого коммутатора *SW*1. Интегрированный сетевой интерфейс *PN*002.*i*01 узла *PNode*002 с функцией сетевой загрузки подсоединен входом/выходом к порту 03 управляемого коммутатора *SW*1. Интегрированный сетевой интерфейс *PN*003.*i*01 узла *PNode*003 с функцией сетевой загрузки подсоединен входом/выходом к порту 04 управляемого коммутатора *SW*1. Интегрированный сетевой интерфейс *PN*004.*i*01 узла *PNode*004 с функцией сетевой загрузки подсоединен входом/выходом к порту 05 управляемого коммутатора *SW*1. Интегрированный сетевой интерфейс *PN*005.*i*01 узла *PNode*005 с функцией сетевой загрузки подсоединен входом/выходом к порту 06 управляемого коммутатора *SW*1. Мастер – узел *MNode*001 соединен входом/выходом двунаправленным внешним сетевым интерфейсом *Gi*_001,1 к порту 12 управляемого коммутатора *SW*1. Мастер – узел *MNode*001 соединен

входом/выходом двунаправленным внешним сетевым интерфейсом *Gi*_001,2 с сетью *Ethernet*. Порт 11 управляемого коммутатора *SW*1 соединен с портом 12 сетевого интерфейса управляемого коммутатора *SW*2, а порт 10 управляемого коммутатора *SW*1соеденен с портом 12 управляемого коммутатора *SW*3 и образуют подсеть *vp*123 для *Web* конфигурирования и диагностики коммутаторов.

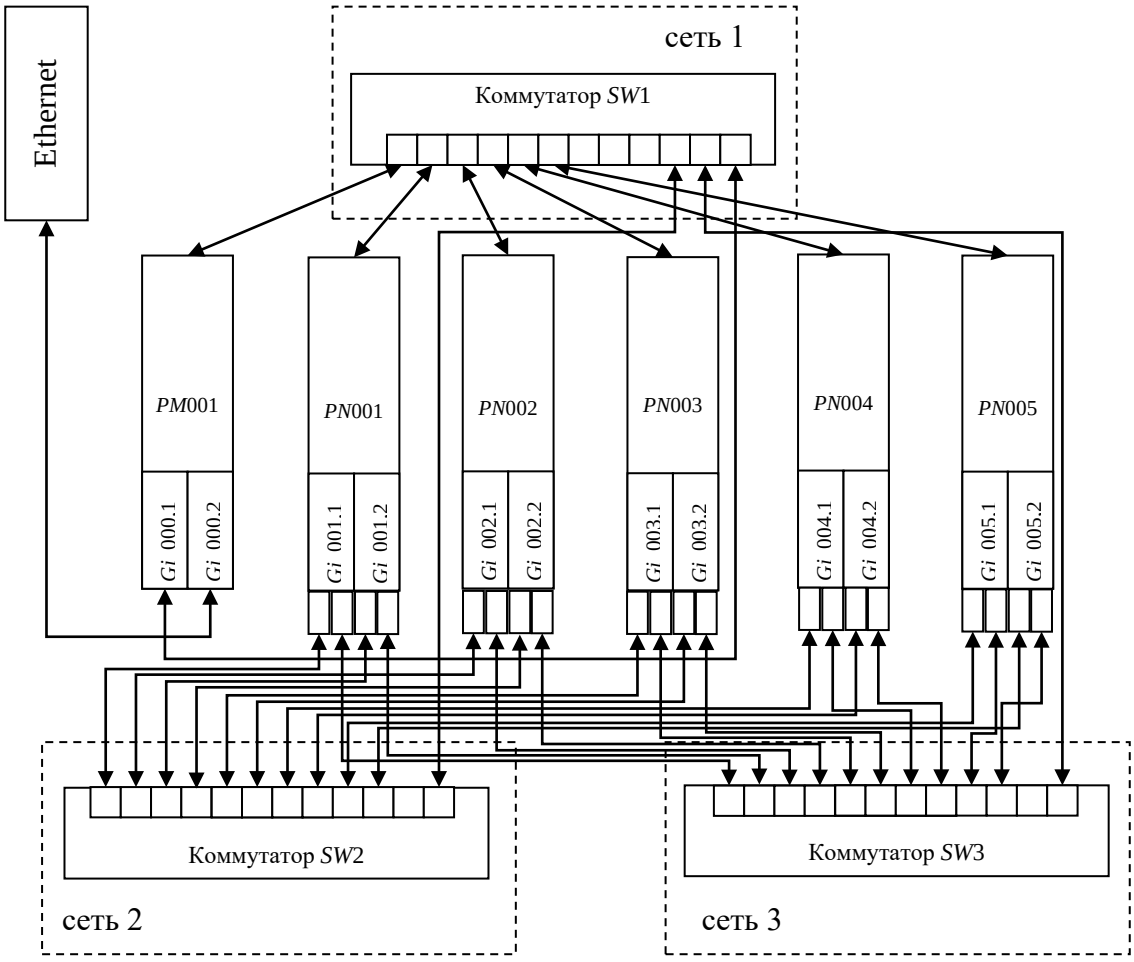

Рис. 1.10. Схема организации сетевого интерфейса по технологии channel bonding

Особенности организации сетевого интерфейса для режима агрегирования каналов (рис.1.10) состоят в том, что в сети обмена данными многопроцессорной вычислительной системы сконфигурированы две симметрично работающие вычислительные сети (сеть 2 и сеть 3) на основе двух коммутационных матриц (коммутаторы *SW* 2 и *SW* 3). Архитектура

вычислительной сети кластера для реализации режима *channel bonding* организована следующим образом.

Slave – узел *PNode*001 соединен входом/выходом двухпортовым двунаправленным внешним сетевым интерфейсом *Gi*_001,1 портом 1 с портом 01 управляемого коммутатора *SW*2, а портом 2 с портом 01 управляемого коммутатора *SW*3. Кроме того, дополнительно такой узел соединен двухпортовым двунаправленным сетевым интерфейсом *Gi*_001,2 портом 1 с портом 02 управляемого коммутатора *SW*2, а портом 2 с портом 02 управляемого коммутатора *SW*3.

Slave – узел *PNode*002 соединен входом/выходом двухпортовым двунаправленным внешним сетевым интерфейсом *Gi*_002,1 портом 1 с портом 03 управляемого коммутатора *SW*2, а портом 2 с портом 03 управляемого коммутатора *SW*3. Кроме того, дополнительно такой узел соединен двухпортовым двунаправленным сетевым интерфейсом *Gi*_002,2 портом 1 с портом 04 управляемого коммутатора *SW*2, а портом 2 с портом 04 управляемого коммутатора *SW*3.

Slave – узел *PNode*003 соединен входом/выходом двухпортовым двунаправленным внешним сетевым интерфейсом *Gi*_003,1 портом 1 с портом 05 управляемого коммутатора *SW*2, а портом 2 с портом 05 управляемого коммутатора *SW*3. Кроме того, дополнительно такой узел соединен двухпортовым двунаправленным сетевым интерфейсом *Gi*_003,2 портом 1 с портом 06 управляемого коммутатора *SW*2, а портом 2 с портом 06 управляемого коммутатора *SW*3.

Slave – узел *PNode*004 соединен входом/выходом двух портовым двунаправленным внешним сетевым интерфейсом *Gi*_004,1 портом 1 с портом 07 управляемого коммутатора *SW*2, а портом 2 с портом 07 управляемого коммутатора *SW*3. Кроме того, дополнительно такой узел соединен двухпортовым двунаправленным сетевым интерфейсом *Gi*_004,2 портом 1 с портом 08 управляемого коммутатора *SW*2, а портом 2 с портом 08 управляемого коммутатора *SW*3.

Slave – узел *PNode*005 соединен входом/выходом двухпортовым двунаправленным внешним сетевым интерфейсом *Gi*_005,1 портом 1 с портом 09 управляемого коммутатора *SW*2, а портом 2 с портом 09 управляемого коммутатора *SW*3. Кроме того, дополнительно такой узел соединен двухпортовым двунаправленным сетевым интерфейсом *Gi*_005,2 портом 1 с портом 10 управляемого коммутатора *SW*2, а портом 2 с портом 10 управляемого коммутатора *SW*3.

Приведенная схема организации сетевого интерфейса включает на каждом вычислительном узле по две однотипные двух портовые сетевые карты и два однотипных коммутатора. Для конфигурации приведенных сетевых интерфейсов выполняются основные операции по настройке режима *Link Aggregation*.

Как и в ранее рассмотренной многопроцессорной кластерной системе обмен данными осуществляется между соседними узлами. В таком случае связь между узлами кластера организуется по топологии кольцо (рис. 1.11), т.е. узел *PN*001 обменивается данными с *PN*002, узел *PN*002 с *PN*003, узел *PN*003 с *PN*004, узел *PN*004 с *PN*005, узел *PN*005 с *PN*001. За счет реализации режима агрегации каналов скорость обмена данными в сети возрастает в два раза.

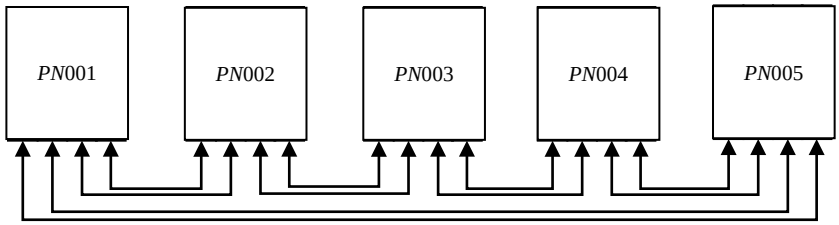

Рис. 1.11. Структура вычислительной сети кластера для реализации граничного обмена в режиме агрегации каналов

Предложенный подход позволяет организовать равномерное распределение нагрузки между соответствующими узлами кластерной системы, а также увеличить скорость обмена данными между узлами кластерной системы. Очевидно, что чем выше будет пропускная способность сети, тем

быстрее будут решаться параллельные задачи, обрабатываемые при помощи модульной кластерной системы.

1.7.2. Особенности подбора элементов сетевого интерфейса кластерной многопроцессорной системы для режима агрегации каналов сетевого интерфейса

Для освещения вопросов, связанных с подбором элементов сетевого интерфейса кластерной многопроцессорной системы, укажем некоторые особенности их функционирования в многосвязанном режиме.

Вообще отметим, что агрегация каналов (англ. *Link aggregation, trunking*) или стандарт *IEEE* 802.3*ad* – это технология объединения нескольких физических каналов в один логический. Реализация такого подхода способствует не только значительному увеличению пропускной способности магистральных каналов коммутатор – коммутатор или коммутатор – сервер, но и повышению их надежности. При этом заметим, что хотя уже существует стандарт *IEEE* 802.3*ad*, многие компании еще используют для своих продуктов патентованные или вообще закрытые технологии.

Главное преимущество агрегации каналов состоит в том, что радикально повышается скорость обмена данными. При этом основная особенность сетевых адаптеров состоит в следующем. В случае отказа адаптера трафик посылается следующему работающему адаптеру без прерывания сервиса. Если же адаптер вновь начинает работать, то через него опять пересылаются данные.

Рассмотрим сам принцип связывания в параллель нескольких *Ethernet*-адаптеров. Допустим, есть два адаптера *Ethernet*: *eth*0 и *eth*1. Их необходимо объединить в псевдо-*Ethernet*-адаптер *eth*3. При этом система распознает эти агрегированные адаптеры как один. Все агрегированные адаптеры настраиваются на один *MAC*-адрес, поэтому удаленные серверы обращаются с ними как с одним адаптером. Псевдо-адаптер *eth*3 можно настроить на один *IP* адрес, как любой *Ethernet* адаптер. Из-за этого программы обращаются к нему как к самому обычному адаптеру, скорость которого в два раза выше.

Протоколы агрегирования сетевого интерфейса определяют, какие порты используются для исходящего трафика, или какой конкретный порт принимает входящий трафик. Состояние интерфейса используется для проверки линка, является он активным или нет.

Указанные особенности организации сетевого интерфейса для формирования режима агрегации каналов требуют наличия на каждом вычислительном узле двух двухпортовых однотипных сетевых карт *Intel Server Pro*/1000 *MT Dual* и двух однотипных коммутаторов (3*Com SuperStack* 3 *Switch* 3812).

Для конфигурации приведенных сетевых интерфейсов выполняются основные операции по настройке режима *Link Aggregation*. Более детально тип сетевого оборудования и его ориентировочная цена на сегодняшний день приведены в табл. 1.9.

Таблица 1.9. Технические характеристики сетевого оборудования кластерной системы в режиме агрегации каналов сети

Сетевой кабель	Тип	*GigabitEthernet*
	Пропускная способность	1 Гбит/с
	Стандарт	5*e*
	Цена	$ 6
Сетевой адаптер	Тип	*PWLA*8492*MT*
	Производитель	*Intel Server Pro*/1000 *MT Dual*
	Пропускная способность	2x1 Гбит/с
	Цена	$200
Коммутатор	Тип	3*Com SuperStack* 3 *Switch* 3812
	Производитель	3*COM*
	Пропускная способность	24 Гб/с
	Цена	$ 940

1.7.3. Особенности организации и настройки многоканального сетевого интерфейса многопроцессорной системы

Рассмотрим некоторые важные особенности организация сетевого интерфейса в соответствии с технологией *channel bonding*. На первом этапе укажем требования, предъявляемые к аппаратным и программным средствам многопроцессорной кластерной системы. Так, все лезвия многопроцессорной системы должны иметь одинаковый набор *bonded networks*, т.е. нельзя на одном лезвии использовать сетевую карту типа 2*x*100*BaseTx*, а на другом – 10*Base* и 100*BaseTx*. Режим работы сетевых карт тоже должен быть единообразным. Другими словами, недопустим вариант, когда одна сетевая карта функционирует в режиме *full duplex*, а другая в полудуплексном режиме. Технология *channel bonding* требует наличия, как минимум, двух физических подсетей. Но, при необходимости, связанный канал можно построить на основе трех или более сетевых карт.

Для связывания сетевых карт в один канал (одну виртуальную карту) необходимо либо скомпилировать ядро ОС с поддержкой режима *channel bonding*, либо загрузить в ОС модуль ядра *bonding.o*.

Заметим, что в ОС *Linux*, начиная с ядер 2.4.*x*, технология *channel bonding* является стандартной включаемой опцией. Например, в дистрибутиве *Alt Linux Master* 2.2 технология *channel bonding* поставляется в виде загружаемого модуля ядра.

Для конфигурации связанного канала требуется стандартная команда *ifconfig* и, возможно, дополнительная команда *ifenslave*. Это объясняется тем, что программа *ifenslave* позволяет копировать установки первого интерфейса на все остальные дополнительные интерфейсы.

Для рассматриваемой многопроцессорной системы был реализован режим формирования двух подсетей. В этой связи, процесс настройки технологии *channel bonding* в настоящей монографии будет приведен на примере использования двух сетевых карт. Сетевой интерфейс для первой карты должен быть заранее сконфигурирован и полностью работоспособен. Для

добавления в систему второй сетевой карты и объединения ее с первой в связанный канал требуется выполнить некоторые операции. Предварительно останавливаются сетевые интерфейсы в многопроцессорной системе при помощи команды

```
/etc/rc.d/init.d/network stop
```

После этого переходят собственно к конфигурации связанного канала. На первом этапе необходимо изменить файл */etc/modules.conf*, добавив в него следующую строку:

```
alias bond0 bonding
```

Такое добавление сообщает системе о том, что необходимо загрузить модуль *bonding.o*, который определяется также по алиасу *bond*0. Чтобы не перезагружать систему, вручную загружается модуль

```
modprobe bonding
```

Далее переходят в каталог */etc/sysconfig/network-scripts* и переименовывают файл описания первого интерфейса *ifcfg-eth0* в *ifcfg-bond0*:

```
cp ifcfg-eth0 ifcfg-bond0
```

Сформированный файл *ifcfg-bond0* необходимо отредактировать так, чтобы он принял следующий вид:

```
DEVICE=bond0
IPADDR=192.168.1.*
NETMASK=255.255.255.0
NETWORK=192.168.1.0
BROADCAST=192.168.1.255
ONBOOT=yes
BOOTPROTO=none
USERCTL=no
```

Конечно, пользователь должен указать здесь свои собственные *IP*-адрес, маску, адрес сети и *broadcast.* Здесь же приведена та информация, которая используется для связывания каналов освещаемой многопроцессорной кластерной системы. Следует отметить, что мы здесь не удаляли никакие строки из этого файла, а выполнили изменения в необходимых местах. Таким образом, был отредактирован файл описания виртуального сетевого

интерфейса. Следующим шагом должно быть создание файлов описания для двух реальных физических интерфейсов *eth*0 и *eth*1, в которых указывается, что они входят в состав связанного канала. Файлы *ifcfg-eth*0 и *ifcfg-eth*1 должны иметь содержимое, представленное в табл. 1.10.

Таблица 1.10.

Содержимое файлов описания сетевого интерфейса

файл ifcfg-eth0	*файл ifcfg-eth1*
DEVICE=eth0	DEVICE=eth1
USERCTL=no	USERCTL=no
ONBOOT=yes	ONBOOT=yes
MASTER=bond0	MASTER=bond0
SLAVE=yes	SLAVE=yes
BOOTPROTO=none	BOOTPROTO=none

Далее проводится этап инициализации сетевого интерфейса при помощи команды

```
/etc/rc.d/init.d/network start
```

Если дистрибутив системы не позволяет применять *master/slave* нотификацию, то при конфигурации сетевых интерфейсов придется запускать интерфейс связанного канала вручную, используя следующую последовательность команд:

```
/sbin/ifconfig bond0 192.168.1.* up netmask 255.255.255.0 /sbin/ifenslave bond0
eth0 /sbin/ifenslave bond0 eth1
```

Чтобы каждый раз не выполнять приведенные команды вручную, рекомендуется записать их в какой-нибудь *startup*-скрипт, например, в */etc/rc.d/rc.local*, или заменить ими ту часть скрипта */etc/rc.d/init.d/network*, которая ответственна за инициализацию сетевого интерфейса.

Для ручного запуска сетевого интерфейса мы рекомендуем использовать команду *ifenslave*, которая была разработана в рамках проекта *Beowulf*. Пользователь может ее скомпилировать из исходных кодов, которые представлены непосредственно на сайте проекта *Beowulf* [http://beowulf.org/software/ifenslave.c]. Компиляция этой программы осуществляется следующей командой:

```
gcc -Wall -Wstrict-prototypes -O -I/usr/src/linux/include ifenslave.c -o
ifenslave
```

Полученный скомпилированный файл необходимо скопировать в папку */usr/sbin*.

Если по каким-то причинам необходимо, чтобы все сетевые драйверы были инициализированы до загрузки *bonding*-драйвера, следует добавить строку

```
probeall bond0 eth0 eth1 bonding
```

в файл */etc/modules.conf*. Эта инструкция укажет системе, что в случае инициализации интерфейса *bond*0 утилита *modprobe* должна сначала загрузить драйверы для всех сетевых интерфейсов.

Таким образом, настройка технологии *channel bonding* завершается на этом этапе. Если сетевой интерфейс инициализировался без ошибок, это можно проверить, используя команду *ifconfig*. Запустив ее без параметров, пользователь увидит на экране терминала сообщения следующего характера:

```
[root]# /sbin/ifconfig
bond0     Link encap:Ethernet  HWaddr 00:C0:F0:1F:37:B4
          inet addr:192.168.1.*  Bcast:192.168.1.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MASTER MULTICAST  MTU:1500  Metric:1
          RX packets:7224794 errors:0 dropped:0 overruns:0 frame:0
          TX packets:3286647 errors:1 dropped:0 overruns:1 carrier:0
          collisions:0 txqueuelen:0

eth0      Link encap:Ethernet  HWaddr 00:C0:F0:1F:37:B4
          inet addr:192.168.1.*  Bcast:192.168.1.255  Mask:255.255.255.0
          UP BROADCAST RUNNING SLAVE MULTICAST  MTU:1500  Metric:1
          RX packets:3573025 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1643167 errors:1 dropped:0 overruns:1 carrier:0
          collisions:0 txqueuelen:100
          Interrupt:10 Base address:0x1080

eth1      Link encap:Ethernet  HWaddr 00:C0:F0:1F:37:B4
          inet addr:192.168.1.*  Bcast:192.168.1.255  Mask:255.255.255.0
          UP BROADCAST RUNNING SLAVE MULTICAST  MTU:1500  Metric:1
          RX packets:3651769 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1643480 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:100
          Interrupt:9 Base address:0x1400

lo        Link encap:Local Loopback
                 inet addr:127.0.0.1  Mask:255.0.0.0
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
          RX packets:1110 errors:0 dropped:0 overruns:0 frame:0
          TX packets:1110 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
```

Если на экран терминала выводятся сообщения приведенного характера, то можно отметить, что связанный канал успешно сконфигурирован. Как

видно, *IP*- и *MAC*-адреса всех сетевых интерфейсов в приведенном варианте получились одинаковыми. Чтобы *switch* мог нормально работать с таким каналом, необходимо настроить режим *Link Aggrigation*. Ознакомиться с тем, каким образом реализуется такая процедура, можно в документации коммутатора. Для разных моделей коммутаторов и разных версий их программного обеспечения такая операция может настраиваться по-разному. В этой связи мы опустим здесь вопросы настройки *Link Aggrigation* на коммутаторах.

Иногда встречаются сообщения (http://studentbank.ru/ view.php?id=9653&p=2), что в некоторых случаях после инициализации виртуального сетевого интерфейса дополнительные каналы не могут сразу принимать входящие пакеты. Такая ситуация может возникнуть по той причине, что новый *MAC*-адрес дополнительных каналов физически не прописывается в *EPROM* сетевой карты. В результате при старте модуля свитч не знает, что этот *MAC*-адрес присоединен к более, чем одному порту. Для того чтобы сообщить свитчу правильный набор *MAC*-адресов, достаточно непосредственно после инициализации интерфейса выполнить несколько пингов. После того, как *ICMP*-пакеты пройдут через коммутатор по всем виртуальным каналам, внутренняя таблица коммутатора примет правильный вид, и в дальнейшем проблем с приемом пакетов не будет.

В соответствии с приведенной методикой осуществляется инициализация сетевых интерфейсов и остальных узлов кластерной системы.

1.7.4. Исследование основных сетевых характеристик для режима агрегации каналов сетевого интерфейса многопроцессорной системы

С учетом применения технологии *channel bonding,* на первом этапе исследований необходимо уточнить соотношения (1.3) и (1.4), которые характеризуют основные сетевые характеристики кластерной системы. Коэффициент пропускной способности сети кластера будем определять следующем образом:

$$k_s = \frac{V_p \cdot N \cdot k \cdot d}{V_b \cdot k_m}. \qquad (1.24)$$

Здесь k – количество симметричных вычислительных подсетей, которые работают одновременно за счет реализации технологии *channel bonding*, k_m – количество коммутационных матриц в сети обмена данных.

Коэффициент пропускной способности коммутатора (k_b) уточним аналогично:

$$k_b = \frac{V_b \cdot k_m}{V_p \cdot N \cdot k \cdot d}. \qquad (1.25)$$

Кроме того, для анализа согласования выбранной коммутационной шины с возможностями коммутатора, уточним коэффициент полосы пропускания коммутатора пропускания (c_k), который с учетом режима агрегации каналов сетевого интерфейса будет определяться соотношением вида:

$$c_k = \frac{V_b \cdot k_m}{N}. \qquad (1.26)$$

Исходные данные для изучения рассматриваемого режима работы сетевого интерфейса многопроцессорной системы перечислены в табл. 1.11.

Таблица 1.11. Исходные данные для расчета сетевых характеристик кластерной системы с использованием технологии channel bonding

V_p	1 Гбит/с
V_b	24 Гбит/с
k	2
k_m	2

На первом этапе исследований выведем аналитическое соотношение для определения равновесного числа узлов кластерной системы. Для этой цели приравняем коэффициенты $k_s = k_b$:

$$\frac{V_p \cdot N \cdot k \cdot d}{V_b \cdot k_m} = \frac{V_b \cdot k_m}{V_p \cdot N \cdot k \cdot d}. \qquad (1.27)$$

После некоторых преобразований соотношения (1.27) получаем квадратное уравнение, требуемое значение корня которого будет определяться соотношением вида:

$$N = \frac{V_b \cdot k_m}{V_p \cdot k \cdot d}. \qquad (1.28)$$

Анализ соотношения (1.28) показывает, что равновесное число узлов кластерной системы в режиме агрегации каналов сетевого интерфейса, при равных прочих условиях, зависит еще и от количества симметричных вычислительных подсетей (k), которые работают одновременно за счет реализации технологии *channel bonding*, а также и от количества коммутационных матриц в сети обмена данных (k_m).

С учетом заявленных возможностей сетевого интерфейса (табл. 1.11), на основании соотношения (1.28) определим равновесное число узлов кластерной системы. При этом равновесное число узлов кластерной системы соответствует $N = 12$.

Далее, для уточнения особенностей функционирования сетевого интерфейса кластерной системы на основании соотношений (1.24) – (1.26) выполним процедуру моделирования основных его числовых характеристик.

Полученные результаты моделирования сведены в табл. 1.12.

Итак, появились предпосылки для общего анализа полученных результатов. Очевидно, что представленный режим работы, при равных прочих условиях, за счет изменения архитектуры сетевого интерфейса многопроцессорной системы позволяет расширять только полосу пропускания коммутационной шины, оставляя равновесное число узлов кластера неизменным. *Последнее обстоятельство означает, что сформированный режим работы сетевого интерфейса кластерной системы будет предоставлять более широкие возможности для реализации процедуры обмена данными между вычислительными узлами, существенно улучшая характеристики эффективности, быстродействия и надежности функционирования системы.*

Таблица 1.12. Результаты расчета основных сетевых коэффициентов кластера с использованием технологии channel bonding

Колич. узлов, N	k_s	k_b	c_k
1,00	0,08	12,00	48,00
2,00	0,17	6,00	24,00
3,00	0,25	4,00	16,00
4,00	0,33	3,00	12,00
5,00	0,42	2,40	9,60
6,00	0,50	2,00	8,00
7,00	0,58	1,71	6,86
8,00	0,67	1,50	6,00
9,00	0,75	1,33	5,33
10,00	0,83	1,20	4,80
11,00	0,92	1,09	4,36
12,00	1,00	1,00	4,00
13,00	1,08	0,92	3,69
14,00	1,17	0,86	3,43
15,00	1,25	0,80	3,20
16,00	1,33	0,75	3,00
17,00	1,42	0,71	2,82
18,00	1,50	0,67	2,67
19,00	1,58	0,63	2,53
20,00	1,67	0,60	2,40
21,00	1,75	0,57	2,29
22,00	1,83	0,55	2,18
23,00	1,92	0,52	2,09
24,00	2,00	0,50	2,00
25,00	2,08	0,48	1,92

В дальнейшем исследования будут направлены на выявление аналитических зависимостей и числовых характеристик эффективности и ускорения вычислений кластерной системы за счет расширения возможностей сетевого интерфейса. Для этой цели на последующем этапе исследований будут рассмотрены особенности взаимодействия сетевого интерфейса кластерной системы с ее узлами.

Тогда возникают предпосылки для переоценки характеристик ускорения и эффективности кластерной системы. На первом этапе уточним такую характеристику, как коэффициент пропускной способности кластерной системы (k_k). Такой коэффициент будет соответствовать единице для режима дефицита сетевого интерфейса. Заметим, что здесь, как и ранее, равновесное

число узлов кластерной системы соответствует N = 12. Для режима профицита сетевого интерфейса такой коэффициент будет убывать по нелинейной зависимости (рис. 1.5).

Тогда общая оценка пропускной способности сети кластера будет определяться следующим соотношением:

$$V = k \cdot d \cdot V_p \cdot k_k \,. \qquad (1.29)$$

Очевидно, что с учетом выражения (1.29), такая формула будет распадаться на две: одна описывает общую оценку пропускной способности сети кластера для режима дефицита сетевого интерфейса (V_1), а другая – режима его профицита (V_2). Для режима дефицита сетевого интерфейса получим:

$$V_1 = k \cdot d \cdot V_p, \qquad (1.30)$$

для режима профицита сетевого интерфейса такая скорость будет выражаться формулой вида:

$$V_2 = \frac{V_b \cdot k_m \cdot d}{N}. \qquad (1.31)$$

Анализ соотношения (1.30) показывает, что для режима дефицита сетевого интерфейса общая оценка пропускной способности сети кластера будет зависеть от скорости порта узла многопроцессорной сети, количества подсетей и режима передачи данных в вычислительной сети (дуплекс или полудуплекс). В то же время такая скорость не будет зависеть от числа узлов многопроцессорной системы. Данный факт можно объяснить тем, что система работает в режиме дефицита сетевого интерфейса, а это означает, что до равновесного числа узлов кластерной системы скорость коммутации данных в сети будет определяться, в основном, скоростью узла порта многопроцессорной системы.

С другой стороны, в режиме профицита сетевого интерфейса оценка пропускной способности сети кластера (1.31) будет зависеть от пропускной

способности используемого коммутатора, количества коммутационных матриц в сети обмена данных, режима передачи данных в вычислительной сети (дуплекс или полудуплекс) и количества вычислительных узлов кластерной системы. В то же время такая скорость не будет зависеть от количества подсетей сетевого интерфейса. Такой результат очевиден, поскольку коммутационная шина коммутатора может выделить каждому узлу многопроцессорной системы лишь полосу пропускания, которая определяется пропускной способностью коммутаторов и числом коммутационных матриц.

С учетом выявленных обстоятельств и пересмотрим оценки эффективности многопроцессорной кластерной системы.

1.7.5. Исследование оценок эффективности кластерной системы для режима агрегации каналов сетевого интерфейса

Применение технологии *channel bonding* обеспечивает не только повышение скорости обмена данными между узлами кластерной системы, но и снижение загрузки канала, который соединяет узлы кластера. Очевидно, что такая реализация механизма *channel bonding* кластерной вычислительной системы будет повышать ее эффективность. Последующие исследования и направлены на изучение оценок эффективности распараллеливания путем введения дополнительных вычислительных сетей.

Исходные данные для изучения упомянутого режима работы многопроцессорной системы перечислены в табл. 1.13.

Таблица 1.13. 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

На первом этапе определим время граничного обмена кластерной системы для режимов дефицита и профицита сетевого интерфейса:

$$T_{ex1} = \frac{m \cdot (N-1) \cdot \sqrt{R}}{k \cdot d \cdot V_p}, \tag{1.32}$$

$$T_{ex2} = \frac{m \cdot (N-1) \cdot \sqrt{R} \cdot N}{k_m \cdot d \cdot V_b}. \tag{1.33}$$

Далее, для рассматриваемой кластерной многопроцессорной системы в условиях проводимого эксперимента оценим количество узлов кластерной системы, при котором задача будет решаться наиболее эффективно. При этом заметим, что время счета одной итерации вычислительного процесса складывается из двух слагаемых – времени непосредственного счета на процессорах $T_{calc} = \frac{T_{it}}{N}$ и времени обмена данных между вычислительными узлами кластера T_{ex}, т.е.

$$T_{it} = T_{calc} + T_{ex}. \tag{1.34}$$

При этом в [3, 19] показано, что скорость вычислений будет расти примерно до момента, когда

$$T_{calc} \approx T_{ex}. \tag{1.35}$$

Таким образом, на основании соотношения (1.35) можно оценить количество узлов кластерной вычислительной системы, при котором задача будет решаться наиболее эффективно. Заметим, что данный этап исследований имеет своей целью уменьшение общего времени счета путем распараллеливания программы. Очевидно, что при этом общий размер разностной сетки не зависит от числа вычислительных узлов кластерной системы. Учитывая соотношение (1.35), получают аналитические выражения для определения оптимального числа узлов кластерной системы:

$$\frac{T_{it}}{N} \approx \frac{m \cdot (N-1) \cdot \sqrt{R}}{k \cdot d \cdot V_p} \tag{1. 36}$$

для работы кластера в режиме дефицита сетевого интерфейса и

$$\frac{T_{it}}{N} \approx \frac{m \cdot (N-1) \cdot \sqrt{R} \cdot N}{k_m \cdot d \cdot V_b} \qquad (1.37)$$

для работы кластера в режиме профицита сетевого интерфейса. На основании выражений (1.36) и (1.37) можно получить два уравнения относительно N для определения оптимального числа узлов кластерной системы, при котором общее время вычислений, требуемое для решения задачи, будет минимальным.

Уравнение (1.36) преобразуется к квадратичному виду

$$N^2 - N - \frac{T_{it} \cdot k \cdot d \cdot V_p}{m \cdot \sqrt{R}} = 0. \qquad (1.38)$$

Решением такого уравнения будут два корня, при этом один из них положительный, а другой – отрицательный. Исходя из поставленных физических условий задачи, принимается положительный корень, значение которого равно девяти, т.е. $N = 9$. Заметим, что такое решение удовлетворяет неравенству из определения 1.5, которое устанавливает условия функционирования кластерной системы в режиме дефицита сетевого интерфейса.

Уравнение (1.36) сведется к кубическому виду

$$N^3 - N^2 - \frac{T_{it} \cdot k_m \cdot d \cdot V_b}{m \cdot \sqrt{R}} = 0 \qquad (1.39)$$

и будет иметь два мнимых корня и один действительный. Действительный корень соответствует $N = 11$. Однако анализ такого корня показывает, что он не удовлетворяет условию функционирования кластерной системы в режиме профицита сетевого интерфейса (определение 1.6).

На основании полученных соотношений было проведено моделирование основных характеристик эффективности многопроцессорной системы. Полученные результаты сведены в табл. 7.6.

Таблица 1.13. Результаты расчета основных характеристик эффективности при реализации двухканального режима функционирования вычислительной сети кластера

Колич. узлов, N	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. 1.12, 1.13).

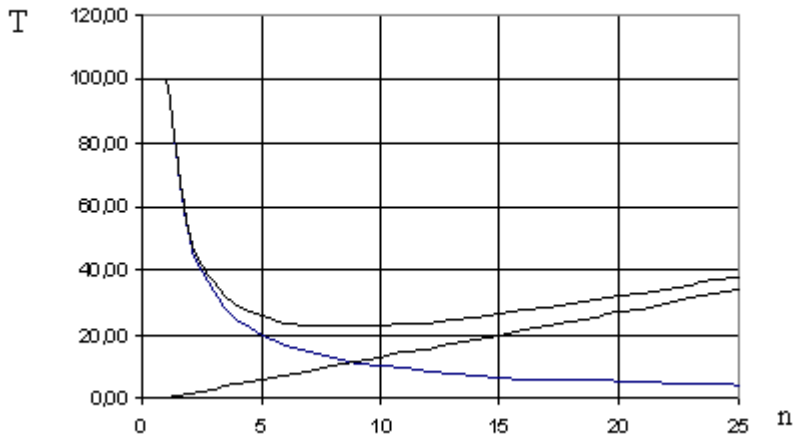

Рис. 1.12. Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы при реализации двухканального режима функционирования вычислительной сети кластера

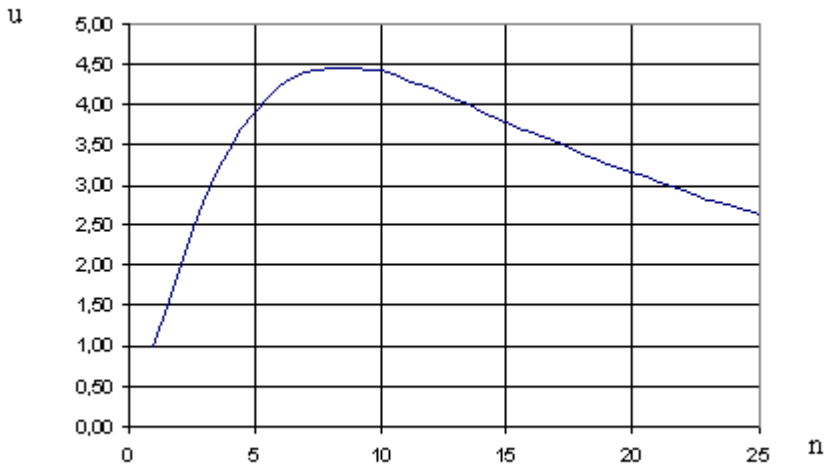

Рис. 1.13. Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы при реализации двухканального режима функционирования вычислительной сети кластера

Итак, имеем предпосылки для количественной оценки эффективности многопроцессорной системы при реализации двухканального режима функционирования вычислительной сети кластера ($к$ = 2). В рамках рассматриваемой задачи оптимальное число узлов кластерной системы, при котором достигается максимальная эффективность распараллеливания, будет соответствовать N = 9. При этом заметим, что цена сетевого оборудования при использовании режима агрегации каналов составила около 5000 у.е.

При выбранном размере кластера задача будет решаться в 4,6 раза быстрее, чем на одном компьютере. Как показывают расчетные данные, такой режим работы кластера позволил не только повысить эффективность системы, но и существенно сократить время вычислений. Так, время вычислений уменьшилось с 30,81 с до 22,4 с. Таких результатов удалось достичь за счет уменьшения времени граничного обмена данными между вычислительными узлами кластерной системы.

1.7.6. Исследование загрузки вычислительной сети кластерной системы для режима агрегации каналов сетевого интерфейса

Рассмотрим характеристику коэффициента использования сети кластерной системы для режима агрегации каналов сетевого интерфейса. Такая

характеристика необходима для проверки правильности подобранного сетевого оборудования. С этой целью выведем соотношение для коэффициента использования сети через параметры кластерной системы для режима дефицита сетевого интерфейса. Выведем коэффициент использования сети кластерной системы через ее параметры.

Принимая во внимание выражения (1.30), (1.32), (1.34), значение коэффициента использования сети можно записать в виде аналитического соотношения, выраженного через параметры вычислительного кластера, т.е.:

$$\xi = \frac{m \cdot N \cdot (N-1) \cdot \sqrt{R}}{T_i \cdot k \cdot d \cdot V_p + N \cdot m \cdot (N-1) \cdot \sqrt{R}}. \qquad (1.40)$$

Результаты расчета коэффициента использования сети для дуплексного режима работы многопроцессорной системы приведены в табл. 1.14.

Таблица 1.14. Результаты расчета коэффициента использования сети кластера для двухканального режима функционирования сетевого интерфейса

Колич. узлов	*КЗС*
1	0,00
2	0,03
3	0,08
4	0,15
5	0,22
6	0,30
7	0,37
8	0,44
9	0,50
10	0,56
11	0,61
12	0,65
13	0,69
14	0,72
15	0,75

Полученные результаты позволяют сделать вывод, что при выбранном режиме функционирования кластера можно использовать не больше девяти

лезвий. Таким образом, можно отметить, что оборудование сетевого интерфейса кластерной системы подобрано удачно.

Раздел 2.
ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ СОВРЕМЕННЫХ КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ И ИССЛЕДОВАНИЕ ИХ ВЛИЯНИЯ НА ЭФФЕКТИВНОСТЬ МНОГОПРОЦЕССОРНЫХ КЛАСТЕРНЫХ СИСТЕМ

В настоящее время существует много различных вариантов построения кластерных вычислительных систем. Однако одно из основных различий в их конструировании лежит в области используемой сетевой технологии, выбор которой определяется, прежде всего, классом решаемых задач. Вычислительная сеть модульной кластерной системы – это модульная и адаптируемая коммутационная система, которая настраивается в соответствии с самыми различными требованиями. Ее модульность облегчает добавление новых компонентов или перемещение существующих, а адаптивность упрощает внесение изменений и усовершенствований.

В этой связи, предварительно отметим, что вычислительная сеть кластерной системы имеет две основные характеристики – пропускную способность и латентность. Пропускная способность вычислительной сети – это скорость передачи данных между двумя ее узлами после того, как связь между ними установлена. Латентность – это среднее время между вызовом функции передачи данных и самой передачей. Такое время затрачивается на адресацию информации, срабатывание промежуточных сетевых устройств, а также другие сетевые ситуации, возникающие при передаче данных.

Вообще заметим, что пропускная способность и латентность не только характеризуют кластер, но и ограничивают класс задач, который может эффективно обрабатываться при помощи кластерной системы. Так, если задача требует интенсивного обмена данных для пакетов небольшой длины, кластер, использующий сетевое оборудование с большой латентностью, будет большую часть времени тратить на установление сетевых соединений, а не на передачу данных между узлами системы. Следовательно, узлы в кластерной системе

будут простаивать, и при таких условиях нельзя добиться значительного увеличения эффективности распараллеливания.

С другой стороны, если пересылаются большие пакеты данных, то влияние периода латентности на эффективность кластера может снижаться за счет того, что передача занимает гораздо больше времени, чем установление самого соединения. Как видно, неудачная реализация сетевого интерфейса порой может свести на нет эффект от увеличения числа используемых процессоров.

Очевидно, новый качественный этап развития многопроцессорных кластерных систем лежит в области использования новых современных сетевых технологий.

Принимая во внимание отмеченное, данный раздел монографии посвящен проблеме исследования перспектив применения современных коммуникационных технологий в многопроцессорных кластерных системах. Основное внимание уделяется влиянию сетевого интерфейса на оценки эффективности кластерной системы. Раскрыты вопросы согласования сетевого интерфейса с вычислительными узлами в многопроцессорной системе.

2.1. Сетевая технология *Myrinet*

На первом этапе исследований рассмотрим особенности формирования архитектуры сетевого интерфейса кластерной системы на основе применения технологии *Myrinet,* предлагаемой компанией *Myricom* [22]. Технология *Myrinet* – это широко применяемый для построения кластерных многопроцессорных систем тип коммуникационной среды. Такая технология наиболее выгодна по соотношению цена/производительность. Кроме того, *Myrinet* является экономически выгодной высокоскоростной сетью, используемой для коммуникаций и коммутации внутри параллельных суперкомпьютеров.

В списке *Top*500 самых мощных компьютеров мира до 28 % кластерных установок (июнь 2005 г.) были построены на основе применения технологии

Myrinet. В 2009 году такой показатель снизился до 2 %. На основе этой коммуникационной технологии построен высокопроизводительный вычислительный кластер ВЦ РАН. Архитектура такого кластера, конфигурация вычислительных узлов, тестирование вычислительной коммуникационной среды детально изложены в [23].

В силу снижения интереса к такой сетевой технологии в настоящей монографии она не получила детального анализа. Тем не менее, заметим, что для малобюджетных кластеров сетевая технология *Myrinet* может оказаться предпочтительной. Кроме того, такой сетевой интерфейс может оказаться полезным, если решаемый класс задач требует интенсивной передачи данных с малым размером пакетов. Это объясняется тем, что *Myrinet* обладает малой латентностью. В то же время компания *Myricom* продолжает развивать и совершенствовать сетевой интерфейс, и, в определенных условиях, такая технология может оказаться приоритетной при конструировании многопроцессорных вычислительных систем.

2.2. Сетевая технология *Fibre Channel*

Сетевая технология *Fibre Channel* (*FC*) (англ. *fibre channel* – волоконный канал) представляет собой тип коммуникационной среды для высокоскоростной передачи данных [24]. Благодаря высокой скорости передачи данных, малой задержке и расширяемости она практически не имеет аналогов в своей области. В последние годы область применения сетевой технологии *FC* постепенно перемещается в сегмент конструирования высокопроизводительных вычислительных систем.

Существуют несколько различных топологий подключения устройств на базе коммуникационной среды Fibre Channel. К таким топологиям относятся следующие: "точка-точка" (*point to point*), кольцо с разделяемым доступом (*arbitrated loop*) и коммутируемая связная архитектура (*switched fabric*).

Наиболее простая топология коммуникационной среды *Fibre Channel* – это топология "точка-точка", которая для своей реализации требует сервер,

адаптер Fibre Channel и устройство хранения данных, оснащенное интерфейсом *Fibre Channel*. После установления соединения используется вся доступная полоса пропускания канала. При этом гарантируется, что кадры будут получены в том же порядке, в каком они были посланы. Такой режим работы сетевого интерфейса удобен для аудио- и видеоприложений, например, видеоконференций. Однако такая топология, в силу ряда причин, не нашла применения в многопроцессорных вычислительных системах.

Кольцо с разделяемым доступом – это топология коммуникационной среды *Fibre Channel,* при которой данные передаются по логически замкнутому контуру. В кольце с разделением доступа (*arbitrated loop*) протокол описывает порядок, в котором узел получает разрешение на передачу данных. В данной топологии устройства подключения играют важную роль в организации инфраструктуры кольца, в его работе и управлении им. Команды *Fibre Channel* поддерживают согласование и доступ к кольцу для передачи данных. Кроме того, предоставляются команды для назначения адресов портов кольца с разделением доступа (*arbitrated-loop port addresses – AL-PA*) различным узлам кольца. Каждый узел в управляемом кольце *Fibre Channel* имеет контур для собственного отключения от кольца и сохранения непрерывности кольца в случае ошибки.

Для многопроцессорных вычислительных систем перспективной коммуникационной средой *Fibre Channel* является связная архитектура, которая позволяет назначать адрес различным портам в коммуникационной сети. Такую архитектуру можно построить на основе коммутаторов. На коммутатор возлагается задача по обнаружению каждого устройства в кольце *Fibre Channel* и добавлению характеристик устройства на *SNS (Simple Name Server)*. Заметим, что стоимость коммутаторов технологии *Fibre Channel* стремительно снижается, поэтому применение коммутируемых связных архитектур становится все более предпочтительным. В топологии коммутируемой связной архитектуры *Fibre Channel (Fibre Channel switched-fabric)* каждое устройство имеет логическое подключение к любому другому устройству. Обеспечение

физического подключения устройств по топологии "каждый с каждым" потребовало бы огромных затрат, так как для N устройств необходимо N^2 портов и физических подключений. В реальности каждое устройство подключается к коммутатору, а коммутатор поддерживает логические подключения между всеми своими портами. Коммутатор для технологии *FC* представляет собой высокоскоростное устройство, которое обеспечивает подключение по схеме "каждый с каждым" и обрабатывает несколько одновременных подключений. Кроме того, коммутатор поддерживает такие службы, как *Fabric Login*.

Коммутаторы могут быть подключены каскадно (в виде иерархии) или в виде сети, что позволяет формировать более сложные конфигурации.

Принимая во внимание отмеченное, в дальнейшем при освещении топологии коммуникационной среды *FC* для многопроцессорной кластерной системы, будет рассмотрен вариант ее коммутируемой связной архитектуры.

2.2.1. Особенности выбора элементов сетевого интерфейса многопроцессорной системы

На первом этапе исследований рассмотрим особенности выбора сетевого интерфейса для технологии *FC* при конструировании модульной многопроцессорной кластерной системы, а также выполним процедуру исследования характеристик ускорения вычислений и эффективности распараллеливания в такой кластерной системе.

Сетевые кабели. История применения технологии *FC* представляет особый интерес. Так, для такой технологии физической средой передачи может быть не только оптическое волокно, но и коаксиал, и витая пара, а архитектура представляет собой смесь канальной и сетевой топологии. За последние годы интерфейс *FC* обрел второе дыхание. Произошло это благодаря симбиозу *Ethernet* и *Fibre Channel - FCoE* (*Fibre Channel over Ethernet*).

В последнее время получила широкое распространение технология *FC* с пропускной способностью V_s = 4 Гбит/с. Такой сетевой интерфейс удачно

себя зарекомендовал именно для конструирования модульных многопроцессорных систем.

Заметим, что оптические волокна бывают двух типов – многомодовые *MMF* (*multi mode fiber*) и одномодовые *SMF* (*single mode fiber*). Многомодовое волокно различается профилем сердцевины, широко используются два их стандарта — 62,5/125 и 50/125.

Сегодня рынок *IT* предлагает несколько типов широкополосных, оптимизированных для работы с лазером многомодовых волокон, которые предназначены для использования в высокоскоростных сетевых инсталляциях. Однако при выборе многомодового волокна необходимо принимать во внимание такие ключевые факторы, как полоса пропускания, максимально допустимая протяженность каналов и его стоимость.

Полоса пропускания. Данный параметр характеризует информационную емкость волокна. Обладающее более широкой полосой пропускания волокно позволяет реализовать и более длинные каналы связи, обеспечивает больший запас по суммарному затуханию оптического сигнала и повышает гибкость проектирования кабельных систем.

Максимально допустимая протяженность оптоволоконного канала. Здесь необходимо иметь в виду, что с увеличением скорости передачи данных по волокну дальность передачи снижается. В этой связи, задав протяженность линии связи, и имея представление о потребностях в полосе пропускания, можно определить список типов волокна для выбора.

Вместе с тем, фактором, во многом определяющим окончательный выбор волокна, является его стоимость. В многопроцессорных вычислительных системах данные с высокой скоростью передаются на небольшое расстояние. По этой причине подойдут менее дорогие волокна класса *OM*2 или *OM*2+. Однако, учитывая стоимость этих типов волокон, наиболее экономичным вариантом будет волокно класса *OM*2. Что касается волокна класса *OM*2+, то для данной сетевой архитектуры оно является непомерно дорогим вариантом.

Сетевой адаптер. В качестве сетевых адаптеров необходимо использовать карты, поддерживающие работу в стандарте *FC*. В этой связи, предпочтение было отдано двухпортовому адаптеру *QLA*2462 [25] производителя *QLogic*. Такой адаптер обеспечивает высокие показатели производительности и доступности данных, а также предоставляет набор интеллектуальных сетевых средств, необходимых при создании модульных многопроцессорных вычислительных систем.

Сетевой адаптер *QLA*2462 является высокопроизводительным и наиболее успешным в своем классе. Существенным отличием адаптера *QLA*2462 является наличие полезного набора сетевых средств, обеспечивающих улучшенную защиту данных, маршрутизацию фреймов и возможность управления сетью кластера.

Кроме того, важным обстоятельством является то, что для всех основных ОС доступен единый драйвер, который подходит ко всем 4 Гбит/с моделям *QLogic*. Это обстоятельство существенно упрощает администрирование таких адаптеров в модульных многопроцессорных системах.

Коммутаторы. При выборе коммутирующих устройств, в первую очередь, необходимо учитывать возможность использования технологии связывания каналов *channel bonding*. С другой стороны, такие коммутаторы должны удовлетворять требованиям, предъявляемым к 4 Гб *Fibre Channel* высокопроизводительным сетям. В этой связи, рекомендуется использовать шестнадцатипортовый коммутатор *QLogic SANbox* 5600*Q* [26]. Такой коммутатор обеспечивают естественную масштабируемость, а также высокую производительность на основе объединяющего шасси в расширяемом и легко администрируемом решении. Кроме того, коммутатор *QLogic SANbox* 5600*Q* позволяет наращивать мощность с помощью технологии объединения отдельных коммутаторов в стек. Такой подход обычно можно найти только в дорогих продуктах высшего класса. Объединение в стек снижает стоимость и увеличивает стабильность решений, поскольку предусматривает отдельный расширяемый маршрут для агрегации сетевого трафика между коммутаторами,

который не мешает передаче данных и не занимает обычные порты устройств, а также упрощает администрирование. Таким образом, можно без лишних расходов достичь такого же уровня производительности и удобства обслуживания, как при использовании архитектуры дорогих шасси. Такое свойство предложенного коммутатора является чрезвычайно важным при проектировании модульных многопроцессорных систем.

Технические характеристики сетевого оборудования конструируемой многопроцессорной кластерной системы приведены в табл. 2.1.

Таблица 2.1. Технические характеристики сетевого оборудования кластерной системы

Сетевой кабель	Тип	*Fibre Channel*
	Пропускная способность	4 Гбит/с
	Стандарт	50/125 класса *OM*2
	Цена	$ 50
Сетевой адаптер	Тип	*QLA*2462
	Производитель	*QLogic*
	Пропускная способность	2x4 Гбит/с
	Цена	$ 1900
Коммутатор	Тип	*QLogic SANbox* 5600*Q*
	Производитель	*QLogic*
	Пропускная способность	136 Гб/с
	Цена	$ 4800

2.2.2. Исследование основных сетевых характеристик многопроцессорной системы

При выбранном типе оборудования (табл. 2.1.) на первом этапе были проведены вычислительные эксперименты с целью определения основных сетевых характеристик кластерной системы. Такие исследования были

проведены с целью решения проблемы согласования устройств сетевого интерфейса.

Для этого необходимо сравнивать общую пропускную способность сети кластера (V_s) и пропускную способность коммутатора (V_b). Затем для дальнейшего анализа сетевого интерфейса кластерной системы выполним исследование коэффициента пропускной способности сети кластера (k_s) и коэффициента пропускной способности коммутатора (k_b). За основу проведения соответствующих расчетов примем аналитические соотношения (1.1) – (1.4). При таких обстоятельствах выполним процедуру моделирования указанных коэффициентов в зависимости от количества узлов кластерной системы.

Исходные данные для изучения области изменения коэффициентов сетевого интерфейса многопроцессорной системы перечислены в табл. 2.2.

Таблица 2.2. Исходные данные для расчета сетевых характеристик кластерной системы

V_p	4 Гбит/с
V_b	136 Гбит/с

На первом этапе исследований, в соответствии с выражением (1.7), определим равновесное число узлов кластерной системы.

Так, расчеты показали, что точка сетевого равновесия соответствует числу узлов кластерной системы, равному N = 17. Это означает, что рассматриваемая кластерная система будет предоставлять более широкие возможности для решения поставленной задачи: уменьшения времени расчетов путем увеличения узлов кластерной системы. С другой стороны, можно отметить правильность выбора сетевого оборудования, т.к. равновесное число узлов кластерной системы полностью согласуется с техническими возможностями выбранного коммутатора. На этом этапе исследований можно

отметить, что для рассматриваемого режима работы вычислительной сети коммутатор подобран удачно.

Далее, для уточнения особенностей функционирования сетевого интерфейса кластерной системы, на основании соотношений (1.3) – (5.5) выполним процедуру моделирования основных его числовых характеристик.

Полученные результаты моделирования сведены в табл. 2.3, а их геометрическая интерпретация представлена на рис. 2.1.

Таблица 2.3. Результаты расчета основных сетевых коэффициентов кластерной системы

Колич. узлов, N	k_s	K_b	c_k
1,00	0,06	17,00	136,00
2,00	0,12	8,50	68,00
3,00	0,18	5,67	45,33
4,00	0,24	4,25	34,00
5,00	0,29	3,40	27,20
6,00	0,35	2,83	22,67
7,00	0,41	2,43	19,43
8,00	0,47	2,13	17,00
9,00	0,53	1,89	15,11
10,00	0,59	1,70	13,60
11,00	0,65	1,55	12,36
12,00	0,71	1,42	11,33
13,00	0,76	1,31	10,46
14,00	0,82	1,21	9,71
15,00	0,88	1,13	9,07
16,00	0,94	1,06	8,50
17,00	1,00	1,00	8,00
18,00	1,06	0,94	7,56
19,00	1,12	0,89	7,16
20,00	1,18	0,85	6,80
21,00	1,24	0,81	6,48
22,00	1,29	0,77	6,18
23,00	1,35	0,74	5,91
24,00	1,41	0,71	5,67
25,00	1,47	0,68	5,44

Проведем предварительный анализ полученных результатов. Очевидно, что с увеличением числа узлов кластерной системы будет возрастать пропускная способность сети кластера (V_s, формула 1.1). Тогда изменение

коэффициента пропускной способности сети кластера (k_s, формула 1.3) будет осуществляться по линейному закону (рис. 2.1, линия 2). С другой стороны, увеличение объема данных, пересылаемых между узлами кластера, приведет к тому, что нагрузка на коммутатор будет возрастать, и его коэффициент пропускной способности (k_b,формула 1.4) будет уменьшаться по нелинейному закону (рис. 2.1, линия 1). Результаты проведенного моделирования основных сетевых коэффициентов кластерной системы для технологии *FC* показали, что при выбранном сетевом оборудовании существенно расширяется область дефицита сетевого интерфейса.

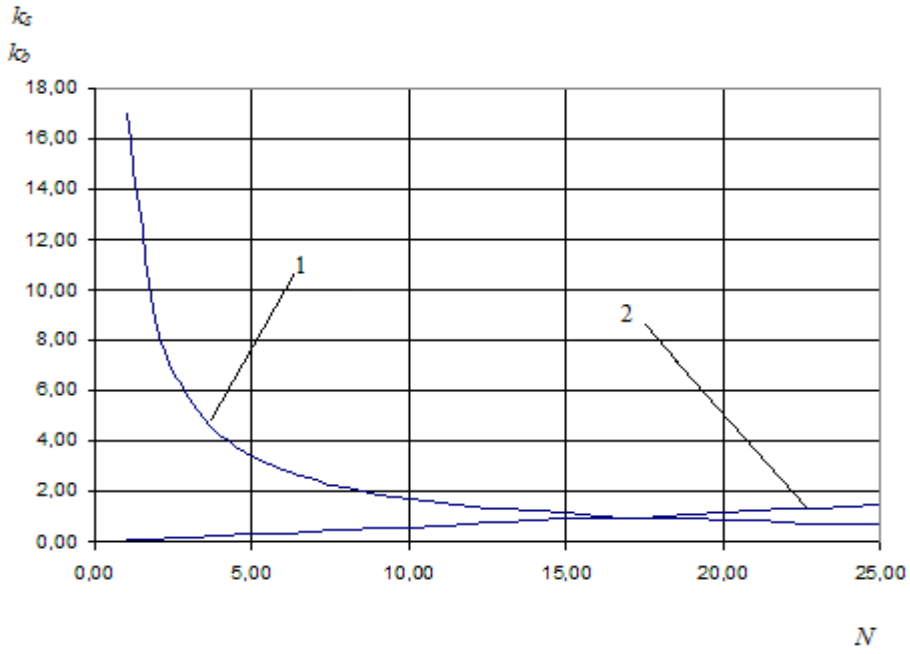

Рис. 2.1. Зависимости основных сетевых коэффициентов кластерной системы от количества узлов

Для более детального анализа предложенного сетевого интерфейса рассмотрим некоторые особенности работы коммутатора. Так, если *N* узлов многопроцессорной системы пытаются установить соединение с одним узлом по протоколу *FC*, то коммутационная шина коммутатора может выделить каждому узлу лишь полосу пропускания (c_k), которая будет определяться соотношением (1.5).

В табл. 2.3 приведен расчет и полосы пропускания коммутатора (c_k). Заметим, что для $N = 17$ значение полосы пропускания на каждый выходной порт кластерной системы будет соответствовать 8 Гбит/с, а это полностью согласуется как с дуплексным режимом обмена данных в многопроцессорной системе, так и с возможностями самой коммутационной шины. При таких обстоятельствах сетевой интерфейс многопроцессорной системы будет функционировать в режиме дефицита и в условиях максимально допустимой загрузки каналов коммутации коммутатора. Таким образом, можно отметить, что технические возможности предложенного коммутатора будут соответствовать протокольным возможностям сети. Следовательно, можно отметить, что сетевое оборудование выбрано удачно.

Итак, проведенный предварительный анализ результатов моделирования основных сетевых коэффициентов кластерной системы создал предпосылки для расчета основных характеристик ее эффективности.

2.2.3. Исследование оценок эффективности кластерной системы

Для характеристики оценок эффективности кластерной системы на первом этапе выполним анализ коэффициента пропускной способности кластерной системы (k_k). Расчет выполнялся на основании соотношений (1.4) и (1.8), а исходные данные соответствовали значениям, принятым в табл. 2.2. График зависимости коэффициента пропускной способности кластерной системы от числа узлов представлен на рис. 2.2.

Анализ такой графической зависимости показывает, что для режима дефицита сетевого интерфейса (когда число узлов кластерной системы $N \leq 17$) коэффициент пропускной способности кластерной системы будет определяться характеристиками сетевого интерфейса. При таких условиях функционирования сетевого интерфейса коммутационная матрица будет работать с наибольшей скоростью.

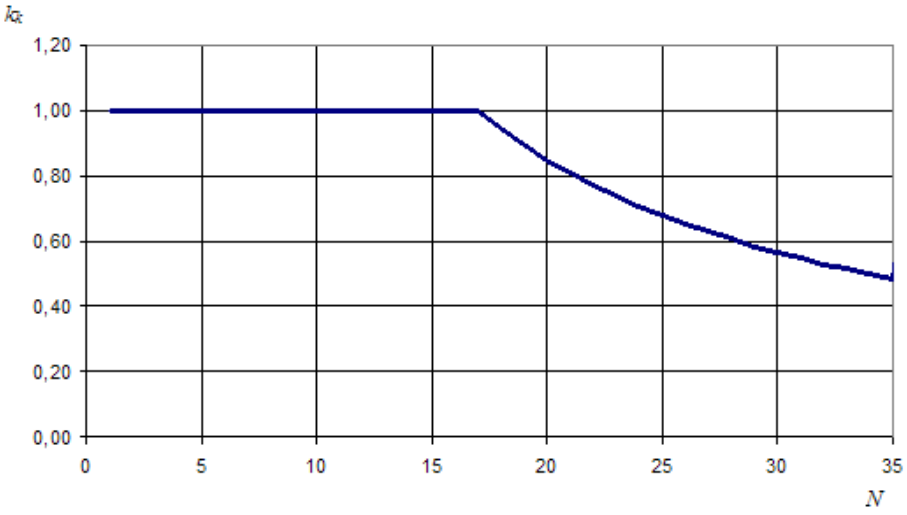

Рис. 2.2. График зависимости коэффициента пропускной способности кластерной системы от числа узлов

Это обстоятельство можно объяснить следующим образом. Механизм управления потоком данных по технологии *FC* требует, чтобы передающий порт не отправлял кадры быстрее, чем принимающий порт может их обработать. Порты *Fibre Channel* имеют буфера для временного хранения кадров и последующей их обработки. Под обработкой подразумевается отправка кадра на другой порт или передача кадра протоколу более высокого уровня. Таким образом, в случае, когда число узлов кластерной системы $N \le 17$, передача данных осуществляется в режиме сквозной коммутации.

В режиме же профицита сетевого интерфейса (когда число узлов кластерной системы $N > 17$) такой параметр будет определяться характеристиками коммутатора, когда сумма входящих пакетов превышает сумму выходящих. Здесь коммутатор переходит в режим коммутации с задержкой фреймов (с использованием буферизации), что приводит к потере его производительности, это обстоятельство и отражено убывающей линией на рис. 2.2.

Исходные данные для исследования оценок эффективности кластерной системы представлены в табл. 2.4.

Таблица 2.4. Исходные данные для расчета характеристик эффективности многопроцессорной системы

V_p	4 Гбит/с
T_{it}	100 с
R	8 Гбит
m	2
d	2

На первом этапе для изучаемой кластерной многопроцессорной системы в условиях рассматриваемого эксперимента оценим количество узлов кластерной системы, при котором задача будет решаться наиболее эффективно. При этом воспользуемся аналитическими соотношениями вида (1.16) – (1.21).

Для работы кластера в режиме дефицита сетевого интерфейса воспользуемся уравнением (1.20). Решением такого уравнения будут два корня, при этом один из них положительный, а другой – отрицательный. Исходя из поставленных физических условий задачи, принимается положительный корень, значение которого равно двенадцати, т.е. $N = 12$. Заметим, что такое решение удовлетворяет неравенству из определения 1.5, которое устанавливает условия функционирования кластерной системы в режиме дефицита сетевого интерфейса.

Для исследования работы кластера в режиме профицита сетевого интерфейса воспользуемся уравнением (1.21), которое будет иметь кубический вид. Решением такого уравнения будет два мнимых корня и один действительный. Действительный корень соответствует $N = 18$. Однако анализ такого корня показывает, что он не удовлетворяет условию функционирования кластерной системы в режиме профицита сетевого интерфейса (определение 1.6).

Отмеченные обстоятельства показывают, что в рамках рассматриваемой задачи оптимальное число узлов кластерной системы, при котором достигается максимальная эффективность распараллеливания, будет соответствовать N =12.

На втором этапе исследований выполним моделирование основных оценок эффективности кластерной системы. Данный этап реализован в соответствии с аналитическими соотношениями, выведенными в работе [10].

Полученные результаты моделирования сведены в табл. 2.5.

Таблица 2.5. Результаты расчета основных характеристик эффективности многопроцессорной системы

Колич. узлов, N	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
16	6,25	10,61	16,86	5,93	0,37
17	5,88	11,31	17,20	5,82	0,34
18	5,56	12,02	17,58	5,69	0,32
19	5,26	12,73	17,99	5,56	0,29

Результаты моделирования представлены также в виде графических зависимостей (рис. 2.3, 2.4).

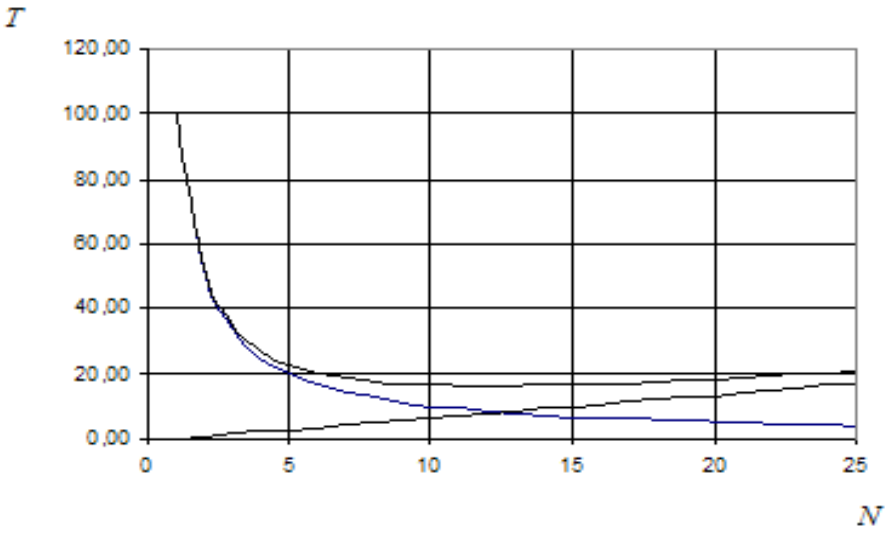

Рис. 2.3. Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы

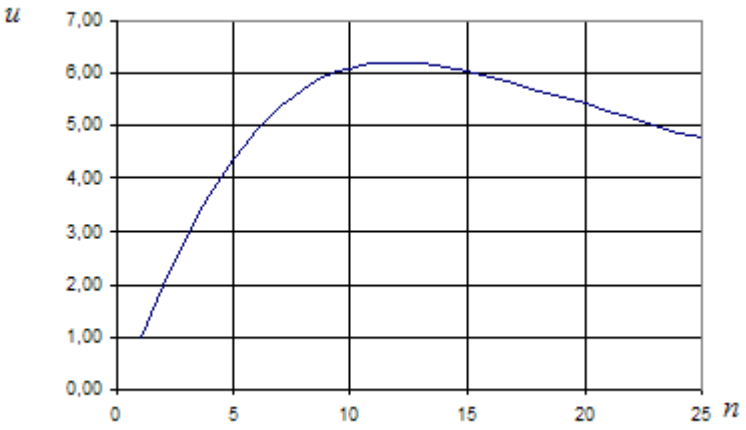

Рис. 2.4. Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы

Проведенный анализ полученных результатов моделирования показал следующее. Оптимальное число узлов кластерной системы, при котором достигается максимальная эффективность распараллеливания, будет соответствовать N =12. Можно отметить, что ориентировочная цена такого сетевого оборудования будет составлять около 34000 у.е.

Наибольшая величина ускорения вычислений при предложенном сетевом интерфейсе соответствует значению, равному 6,21. Время счета задачи уменьшается со 100 с до 16,11 с.

2.2.4. Исследование загрузки вычислительной сети кластерной системы

Для анализа проверки правильности подобранного сетевого оборудования рассмотрим характеристику коэффициента использования сети многопроцессорной кластерной системы. С этой целью воспользуемся аналитическим выражением для расчета коэффициента использования сети, выведенным через параметры кластерной системы (1.23).

Результаты расчета коэффициента использования сети многопроцессорной системы приведены в табл. 2.6.

Таблица 2.6. Результаты расчета коэффициента использования сети кластера

Колич. узлов	*КЗС*
1	0,00
2	0,01
3	0,04
4	0,08
5	0,12
6	0,18
7	0,23
8	0,28
9	0,34
10	0,39
11	0,44
12	0,48
13	0,52
14	0,56
15	0,60
16	0,63
17	0,66
18	0,68
19	0,71
20	0,73

Полученные результаты позволяют сделать вывод, что для оптимального режима сетевого интерфейса необходимо использовать в многопроцессорной системе не более тринадцати лезвий ($N \leq 13$). В данном случае показано, что наилучшими оценки эффективности многопроцессорной системы будут, когда число лезвий многопроцессорной системы равно двенадцати ($N = 12$). Таким образом, можно отметить, что оборудование сетевого интерфейса кластерной системы подобрано удачно.

2.2.5. Высокопроизводительный режим функционирования сетевого интерфейса *FC*

Для повышения надежности функционирования многопроцессорной вычислительной системы, с одной стороны, и улучшения оценок эффективности кластерной системы, с другой стороны, рассмотрим высокопроизводительный режим работы сетевого интерфейса для технологии

FC. В этой связи, на данном этапе исследований рассмотрим особенности применения технологии *FC* с пропускной способностью V_s = 8 Гбит/с. Приведем основные преимущества новой технологии, причем наибольший акцент сделаем не на количественные характеристики, а на качественно новые возможности такой технологии при проектировании многопроцессорных вычислительных систем.

Так, в продуктах *Brocade* 8 Гбит/с появилась новая возможность объединения до трех директоров *DCX* в единую высокопроизводительную фабрику при помощи линков *ICL* (*Inter-Chassis Links*). Это чрезвычайно важное качество нового сетевого интерфейса технологии *FC* для проектирования модульных многопроцессорных систем. Такой подход позволяет в удобной форме реализовать процедуру сопряжения нескольких модулей в единый вычислительный комплекс.

Увеличение пропускной способности каналов позволяет объединять коммутаторы в сети с требуемым уровнем переподписки, но меньшим количеством линков *ISL*. Особенно это актуально для многопроцессорных вычислительных систем, в которых используются внутренние коммутаторы-лезвия. В многопроцессорных вычислительных системах переход части инфраструктуры на 8 Гбит/с значительно упростит работу с кабельной системой в целом (то, что в английской терминологии называется *cable management*), а также процедуры поиска и устранения неисправностей в *SAN*.

Наконец, количество задач, требующих широкую полосу пропускания каналов передачи данных, неуклонно растет. Усиливающаяся популярность технологий виртуализации, увеличение многопоточности приложений, использование больших объемов памяти *RAM*, производительных шин и накопителей позволяют обрабатывать и передавать огромные потоки данных, что, в свою очередь, требует от вычислительной сети все большей и большей производительности.

2.2.5.1. Особенности выбора элементов сетевого интерфейса многопроцессорной системы

Рассмотрим особенности выбора сетевого интерфейса для технологии 8 Гбит/с *FC* при конструировании модульной многопроцессорной кластерной системы, а также выполним процедуру исследования характеристик ускорения вычислений и эффективности распараллеливания в такой кластерной системе.

Сетевые кабели. В рассматриваемой технологии используется многомодовое оптическое волокно (*Multi-Mode Optic*). Здесь целесообразно применять стандарт технологии *FC* 50/125 класса *OM*2. Такой вид кабельной системы может применяться для технологии передачи данных со скоростью до 10 Гбит/с. Производство рассматриваемых кабельных систем отвечает стандарту ISO 9001 и отличается высокой надежностью и долговечностью функционирования.

Сетевой адаптер. В качестве сетевых адаптеров необходимо использовать карты, поддерживающие работу в стандарте *FC*. После детального анализа такого рода адаптеров предпочтение было отдано двухпортовому адаптеру *QLogic QLE*2562 [27]. Такой адаптер обеспечивает высокие показатели производительности и доступности данных, а также предоставляет набор интеллектуальных сетевых средств, необходимых при создании модульных многопроцессорных вычислительных систем.

Сетевой адаптер *QLogic QLE*2562 дает возможность создавать многочисленные логические (виртуальные) соединения, используя один физический порт. Каждое логическое соединение обладает своими собственными ресурсами и может управляться независимо от остальных.

Сетевой адаптер *QLogic QLE*2562 является высокопроизводительным и наиболее успешным в своем классе. Существенным его отличием является наличие полезного набора сетевых средств, обеспечивающих улучшенную защиту данных, маршрутизацию фреймов и возможность управления сетью кластера.

Кроме того, важным обстоятельством является то, что такой адаптер полностью обратно совместим с устройствами ранних версий сетевой технологии *FC* (4 Гбит/с и 2 Гбит/с). Единый драйвер для каждой операционной системы для трех поколений хост-адаптеров *FC* (8 Гбит/с, 4 Гбит/с и 2 Гбит/с) существенно упрощает их внедрение. Унифицированная модель драйвера (прошивка встроена в драйвер) устраняет даже малейшую вероятность несовместимости между версиями драйвера и прошивки.

Коммутаторы. При выборе коммутирующих устройств, в первую очередь, учитывалась не только возможность работы на скорости 8 Гбит/с, но и наличие возможности масштабирования. В этой связи предпочтение было отдано двадцатичетырехпортовому коммутатору *Brocade* 300 [28]. Такой коммутатор обеспечивает естественную масштабируемость, возможность добавления портов по мере роста коммутационной фабрики и высокую производительность на основе объединяющего шасси.

Кроме того, дополнительным свойством коммутатора *Brocade* 300, которое играет важнейшую роль для проектирования модульных многопроцессорных систем, является возможность наращивания его мощности при помощи технологии объединения отдельных коммутаторов в стек. Такое свойство характерно только для дорогих продуктов высшего класса. Объединение в стек снижает стоимость и увеличивает стабильность решений.

Технические характеристики сетевого оборудования конструируемой модульной многопроцессорной кластерной системы приведены в табл. 2.7.

Таблица 2.7. Технические характеристики сетевого оборудования кластерной системы для высокопроизводительного сетевого интерфейса технологии FC

Сетевой кабель	Тип	*Fibre Channel*
	Пропускная способность	8 Гбит/с
	Стандарт	50/125 класса *OM*2
	Цена	$ 30
Сетевой адаптер	Тип	*QLE2562*
	Производитель	*QLogic*
	Пропускная способность	2x8 Гбит/с
	Цена	$ 1900
Коммутатор	Тип	*Brocade* 300
	Производитель	*Brocade*
	Пропускная способность	408 Гб/с
	Цена	$ 6500

2.2.5.2. Исследование основных сетевых характеристик для высокопроизводительного режима функционирования сетевого интерфейса технологии *FC*

С учетом применения высокопроизводительного режима функционирования сетевого интерфейса многопроцессорной кластерной системы, на первом этапе исследований определим основные сетевые характеристики кластерной системы, а именно, коэффициент пропускной способности сети кластера (k_s) и коэффициент пропускной способности коммутатора (k_b). Для этой цели воспользуемся соотношениями (1.1) – (1.4). Кроме того, для анализа согласования выбранной коммутационной шины с возможностями коммутатора, в соответствии с соотношением (1.5), определим коэффициент полосы пропускания коммутатора (c_k).

Исходные данные для изучения рассматриваемого режима работы сетевого интерфейса многопроцессорной системы перечислены в табл. 2.8.

Таблица 2.8. Основные сетевые характеристики кластерной системы с использования высокопродуктивного режима сетевого интерфейса

V_p	8 Гбит/с
V_b	408 Гбит/с

С учетом заявленных возможностей сетевого интерфейса (табл. 2.7), на основании соотношения (1.7), определим равновесное число узлов кластерной системы. При этом расчетное равновесное число узлов кластерной системы соответствует $N = 25$.

Проведем анализ полученных результатов. Очевидно, что представленный режим работы, при равных прочих условиях, за счет изменения архитектуры сетевого интерфейса многопроцессорной системы позволяет не только расширить полосу пропускания коммутационной шины, но и увеличить равновесное число узлов кластера ($N = 25$). *Последнее обстоятельство означает, что сформированный режим работы сетевого интерфейса кластерной системы будет предоставлять более широкие возможности для реализации процесса обмена данными между вычислительными узлами многопроцессорной системы, существенно улучшая характеристики эффективности, быстродействия и надежности функционирования многопроцессорной системы в целом.*

Далее, для уточнения особенностей функционирования сетевого интерфейса кластерной системы, на основании соотношений (1.3) – (1.5), выполним процедуру моделирования основных его числовых характеристик.

Полученные результаты моделирования сведены в табл. 2.9.

Таблица 2.9. Результаты расчета основных сетевых коэффициентов кластера с использованием высокопроизводительного сетевого интерфейса

Колич. узлов, N	k_s	k_b	c_k
1,00	0,04	25,50	408,00
2,00	0,08	12,75	204,00
3,00	0,12	8,50	136,00
4,00	0,16	6,38	102,00
5,00	0,20	5,10	81,60
6,00	0,24	4,25	68,00
7,00	0,27	3,64	58,29
8,00	0,31	3,19	51,00
9,00	0,35	2,83	45,33
10,00	0,39	2,55	40,80
11,00	0,43	2,32	37,09
12,00	0,47	2,13	34,00
13,00	0,51	1,96	31,38
14,00	0,55	1,82	29,14
15,00	0,59	1,70	27,20
16,00	0,63	1,59	25,50
17,00	0,67	1,50	24,00
18,00	0,71	1,42	22,67
19,00	0,75	1,34	21,47
20,00	0,78	1,28	20,40
21,00	0,82	1,21	19,43
22,00	0,86	1,16	18,55
23,00	0,90	1,11	17,74
24,00	0,94	1,06	17,00
25,00	0,98	1,02	16,32
26,00	1,02	0,98	15,69
27,00	1,06	0,94	15,11
28,00	1,10	0,91	14,57
29,00	1,14	0,88	14,07
30,00	1,18	0,85	13,60
31,00	1,22	0,82	13,16
32,00	1,25	0,80	12,75
33,00	1,29	0,77	12,36
34,00	1,33	0,75	12,00
35,00	1,37	0,73	11,66

Результаты расчета основных сетевых коэффициентов кластерной системы с использованием высокопроизводительного сетевого интерфейса на рис. 2.5 представлены и в виде графических зависимостей.

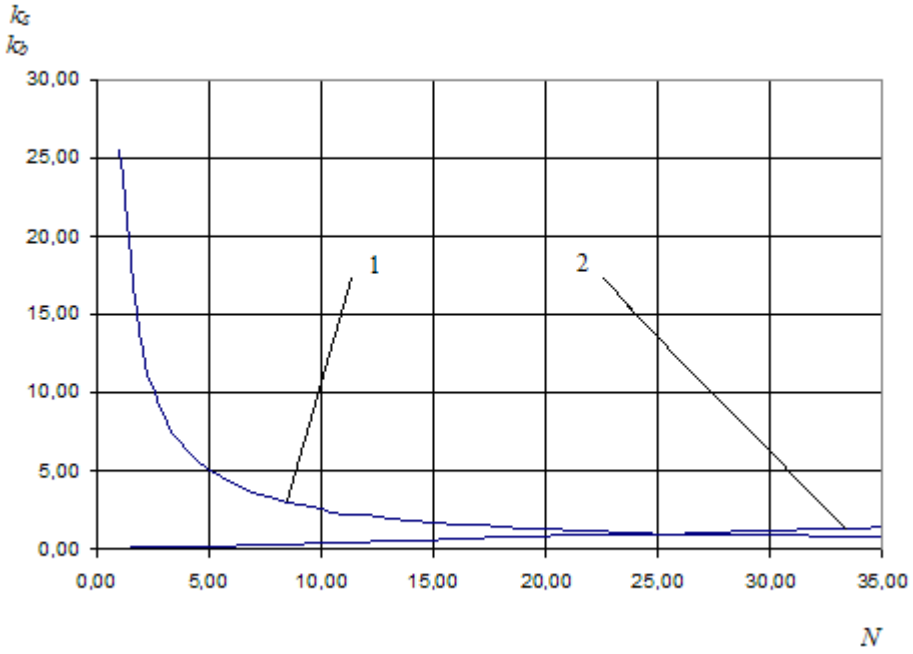

Рис. 2.5. Зависимости основных сетевых коэффициентов кластерной системы от количества узлов с использованием высокопродуктивного режима сетевого интерфейса

Проведем анализ полученных результатов. С увеличением числа узлов кластерной системы изменение коэффициента пропускной способности сети кластера (k_s, формула 1.3) будет осуществляться по линейному закону (рис. 2.5, линия 2). Результаты проведенного моделирования основных сетевых коэффициентов кластерной системы для высокопроизводительного режима работы технологии *FC* показали что, при выбранном сетевом оборудовании существенно расширяется область дефицита сетевого интерфейса. Следовательно, для такой кластерной системы увеличение числа вычислительных узлов будет приводить к существенному улучшению оценок эффективности процесса распараллеливания.

Наконец, в табл. 2.9 приведен расчет и полосы пропускания коммутатора (c_k). Заметим, что для N = 25 значение полосы пропускания на каждый выходной порт кластерной системы будет соответствовать 16 Гбит/с, а это полностью согласуется как с дуплексным режимом обмена данных в многопроцессорной системе, так и с возможностями самой коммутационной шины. При таких обстоятельствах сетевой интерфейс многопроцессорной

системы будет функционировать в режиме дефицита и в условиях максимально допустимой загрузки каналов коммутации коммутатора. Таким образом, можно отметить, что технические возможности предложенного коммутатора будут соответствовать протокольным возможностям сети. Следовательно, можно отметить, что сетевое оборудование выбрано удачно.

Проведенный предварительный анализ результатов моделирования основных сетевых коэффициентов кластерной системы создал предпосылки для расчета основных характеристик ее эффективности.

2.2.5.3. Исследование оценок эффективности кластерной системы в режиме агрегации каналов

В соответствии со сформированным режимом функционирования сетевого интерфейса проведем этап моделирования основных оценок эффективности кластерной системы.

Для всестороннего освещения процессов, протекающих в многопроцессорной вычислительной системе, на первом этапе исследований проведем анализ коэффициента пропускной способности кластерной системы (k_k). Расчет выполнялся на основании соотношений (1.4) и (1.8), а исходные данные соответствовали значениям, принятым в табл. 2.8.

График зависимости коэффициента пропускной способности кластерной системы от числа узлов представлен на рис. 2.6.

Анализ такой графической зависимости показывает, что для режима дефицита сетевого интерфейса (когда число узлов кластерной системы $N \leq 25$) коэффициент пропускной способности кластерной системы будет определяться характеристиками сетевого интерфейса. При таких условиях функционирования сетевого интерфейса коммутационная матрица будет передавать данные с максимально возможной скоростью. Таким образом, в случае, когда число узлов кластерной системы $N \leq 25$, передача данных осуществляется в режиме сквозной коммутации.

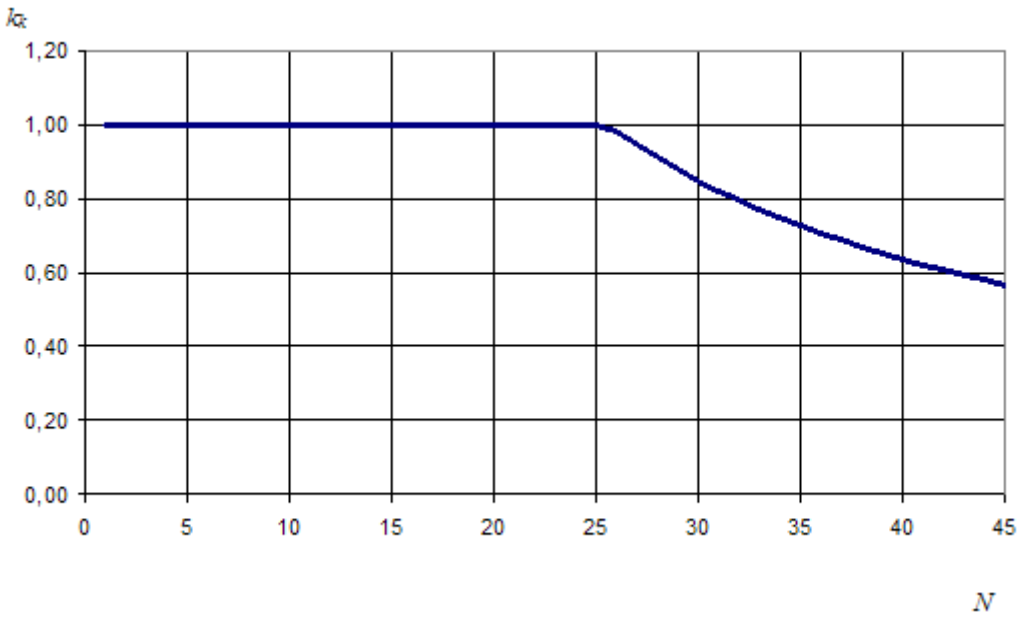

Рис. 2.6. График зависимости коэффициента пропускной способности кластерной системы от числа узлов для высокопроизводительного режима функционирования технологии FC

В режиме же профицита сетевого интерфейса (когда число узлов кластерной системы $N > 25$) такой параметр будет определяться характеристиками коммутатора, когда коммутационная матрица начинает функционировать в режим буферизации с задержкой фреймов. А это приводит к потере его производительности, что и отражено убывающей линией на рис. 2.6.

Заметим, что для предложенного сетевого интерфейса был выбран 24х-портовый коммутатор *Brocade* 300, но, в то же время, равновесное число узлов составляет $N = 25$. Последнее обстоятельство означает, что вычислительная сеть многопроцессорной системы при любом числе узлов (но не более 24) всегда будет функционировать в режиме дефицита сетевого интерфейса. Последнее обстоятельство говорит об удачном подборе комплектующих вычислительной сети кластерной системы.

Заметим, что элементы сетевого интерфейса кластерной системы выбирались с учетом высокопроизводительного режима технологии *FC*. Такой подход позволяет существенно улучшить оценки эффективности модульной многопроцессорной кластерной системы. В таком случае особый интерес будут

представлять оценки эффективности многопроцессорной вычислительной системы. Исходные данные для исследования оценок эффективности такой кластерной системы представлены в табл. 2.10.

Таблица 2.10. Исходные данные для расчета характеристик эффективности многопроцессорной системы для высокопроизводительного режима функционирования технологии FC

V_p	8 Гбит/с
T_{it}	100 с
R	8 Гбит
m	2
d	2

На первом этапе для изучаемой кластерной многопроцессорной системы, в условиях рассматриваемого эксперимента, оценим количество узлов кластерной системы, при котором задача будет решаться наиболее эффективно. При этом воспользуемся аналитическими соотношениями вида (1.16) – (1.21).

Для работы кластера в режиме дефицита сетевого интерфейса воспользуемся уравнением (1.20). Решением такого уравнения будут два корня, при этом один из них положительный, а другой – отрицательный. Исходя из поставленных физических условий задачи, принимается положительный корень, значение которого равно семнадцати, т.е. $N = 17$. Заметим, что такое решение удовлетворяет неравенству из определения 1.5, которое устанавливает условия функционирования кластерной системы в режиме дефицита сетевого интерфейса.

Для исследования работы кластера в режиме профицита сетевого интерфейса воспользуемся уравнением (1.21), которое будет иметь кубический вид. Решением такого уравнения будут два мнимых корня и один действительный. Действительный корень соответствует $N = 24$. Однако анализ такого корня показывает, что он не удовлетворяет условию функционирования кластерной системы в режиме профицита сетевого интерфейса (определение 1.6).

Отмеченные обстоятельства показывают, что в рамках рассматриваемой задачи оптимальное число узлов кластерной системы, при котором достигается максимальная эффективность распараллеливания, будет соответствовать N =17.

На втором этапе исследований выполним моделирование основных оценок эффективности кластерной системы. Данный этап реализован в соответствии с аналитическими соотношениями, выведенными в работе [10].

Полученные результаты моделирования сведены в табл. 2.11.

Таблица 2.11. Результаты расчета основных характеристик эффективности многопроцессорной системы для высокопроизводительного режима функционирования технологии FC

Колич. узлов, N	T_n	T_{ex}	*T*	*USK*	*EF*
1	100,00	0,00	100,00	1,00	1,00
2	50,00	0,35	50,35	1,99	0,99
3	33,33	0,71	34,04	2,94	0,98
4	25,00	1,06	26,06	3,84	0,96
5	20,00	1,41	21,41	4,67	0,93
6	16,67	1,77	18,43	5,42	0,90
7	14,29	2,12	16,41	6,09	0,87
8	12,50	2,47	14,97	6,68	0,83
9	11,11	2,83	13,94	7,17	0,80
10	10,00	3,18	13,18	7,59	0,76
11	9,09	3,54	12,63	7,92	0,72
12	8,33	3,89	12,22	8,18	0,68
13	7,69	4,24	11,93	8,38	0,64
14	7,14	4,60	11,74	8,52	0,61
15	6,67	4,95	11,62	8,61	0,57
16	6,25	5,30	11,55	8,66	0,54
17	5,88	5,66	11,54	8,67	0,51
18	5,56	6,01	11,57	8,65	0,48
19	5,26	6,36	11,63	8,60	0,45
20	5,00	6,72	11,72	8,53	0,43
21	4,76	7,07	11,83	8,45	0,40
22	4,55	7,42	11,97	8,35	0,38
23	4,35	7,78	12,13	8,25	0,36
24	4,17	8,13	12,30	8,13	0,34
25	4,00	8,49	12,49	8,01	0,32

Результаты моделирования представлены также в виде графических зависимостей (рис. 2.7, 2.8).

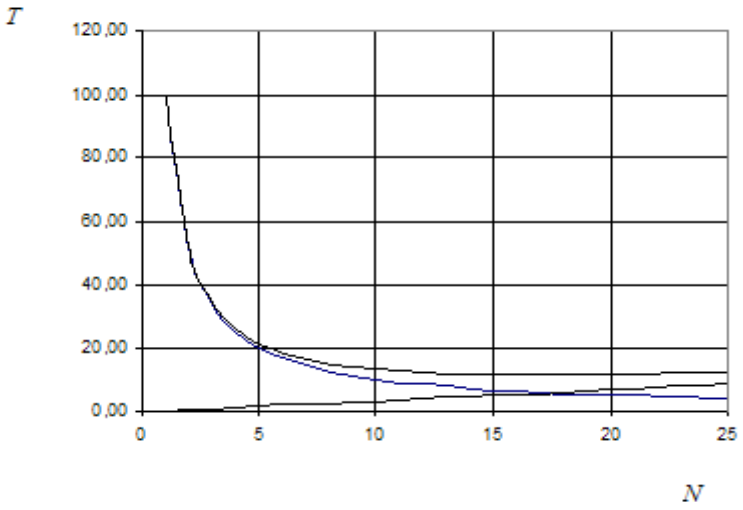

Рис. 2.7. Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы для высокопроизводительного режима функционирования технологии FC

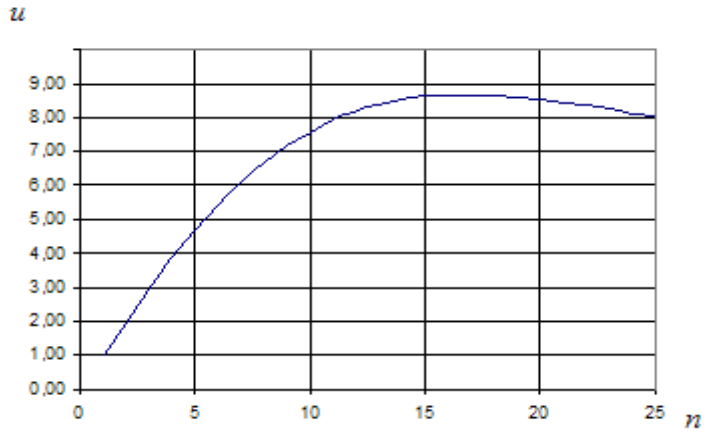

Рис. 2.8. Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы для высокопроизводительного режима функционирования технологии FC

Проведенный анализ полученных результатов моделирования показал следующее. Оптимальное число узлов кластерной системы, при котором достигается максимальная эффективность распараллеливания, соответствует $N = 17$, при этом ориентировочная цена сетевого оборудования будет составлять около 50000 у.е. Наибольшая величина ускорения вычислений при предложенном сетевом интерфейсе соответствует значению, равному 8,67. Время счета задачи уменьшается со 100 с до 11,54 с.

2.2.5.4. Исследование загрузки вычислительной сети кластерной системы в режиме агрегации каналов

Для анализа проверки правильности подобранного сетевого оборудования рассмотрим характеристику коэффициента использования сети многопроцессорной кластерной системы. С этой целью воспользуемся аналитическим выражением, выведенным через параметры кластерной системы (1.23), для расчета коэффициента использования сети.

Результаты расчета коэффициента использования сети многопроцессорной системы приведены в табл. 2.12.

Таблица 2.12. Результаты расчета коэффициента использования сети кластера на основе применения технологии агрегации каналов

Колич. узлов	*КЗС*
1	0,00
2	0,01
3	0,02
4	0,04
5	0,07
6	0,10
7	0,13
8	0,17
9	0,20
10	0,24
11	0,28
12	0,32
13	0,36
14	0,39
15	0,43
16	0,46
17	0,49
18	0,52
19	0,55
20	0,57

Полученные результаты позволяют сделать вывод, что для оптимального режима сетевого интерфейса необходимо использовать в многопроцессорной системе не более восемнадцати лезвий (18). В данном случае показано, что

наилучшими оценки эффективности многопроцессорной системы будут, когда число лезвий многопроцессорной системы равно семнадцати ($N = 17$). Таким образом, можно отметить, что оборудование сетевого интерфейса кластерной системы подобрано удачно.

2.3. Сетевая технология 10*Gb Ethernet*

Сетевая технология 10 *Gigabit Ethernet* или 10*GbE* является новейшим и самым быстрым из существующих стандартов *Ethernet* [29]. Он определяет версию *Ethernet* с номинальной скоростью передачи данных 10 Гбит/с.

Сетевая технология 10 *Gigabit Ethernet* представляет собой четвертое поколение *Ethernet*. После просто *Ethernet, Fast Ethernet* и утвердившейся в своем окончательном варианте *Gigabit Ethernet* уже используется сетевая архитектура *Ethernet* с пропускной способностью 10 Гбит/с.

Еще несколько лет назад казалось, что скорость передачи 10 Гбит/с будет прерогативой исключительно оптических решений, однако прогнозы не оправдались. Стандарт *IEEE* 802.3*an* (10*GBaseT*) на 10 *Gigabit Ethernet* по медному кабелю был принят еще в 2006 г., и многие производители сетевых кабельных систем предлагают продукты для поддержки соответствующего активного оборудования. Существующие стандарты для 10 *Gigabit Ethernet* активно используются многими производителями, и решения на их основе реализуются в реальных проектах. Активное оборудование, основанное на стандартах 10 Гбит/с, выпускают многие производители, например, *Alcatel, Cisco Systems, Enterasys Networks, Extreme Networks, Force10 Networks, Foundry Networks, Hewlett-Packard (HP)* и др. Однако стоит такое оборудование очень дорого. И хотя цены на него падают, в ближайшем будущем предприятиям придется платить десятки тысяч долларов за каждое 10 Гбит/с сетевое соединение. Конечно, вузовским подразделениям, "посаженным" на строгую финансовую диету, непросто выкроить средства из своих скудных бюджетов на покупку данных устройств. В этой связи, при проектировании

многопроцессорных кластерных систем речь должна идти об экономической целесообразности таких проектов.

Принимая во внимание отмеченное, в данной монографии и будет исследован сетевой интерфейс на базе технологии *10Gb Ethernet.*

Вообще заметим, что в настоящее время активно ведется разработка стандартов интерфейсов 40 и 100 Гбит/с для медной проводки. Конечно, пока трудно даже прогнозировать, сколько будут стоить такие решения, но, несомненно, они найдут свою сферу применения.

2.3.1. Особенности выбора элементов сетевого интерфейса

Рассмотрим особенности выбора сетевого интерфейса для технологии 10*Gb Ethernet* при конструировании модульной многопроцессорной кластерной системы.

Сетевые кабели. Итак, рассматривается новое поколение кабельных систем стандарта 10*GBASE-T* [30]. Целью разработки стандарта 10*GBase-T* является создание недорогой стандартизированной 10-гигабитовой технологии передачи данных по медному кабелю из витых пар. Технология 10*GBase-T* TIA/EIA-568-B для кабельных систем категории 6 (*Cat* 6) имеет большое значение для построения на ее основе многопроцессорных вычислительных систем. Внедрение такой технологии должно привести к снижению стоимости соединений 10-*Gigabit Ethernet* на 50 – 80%, а последующие поколения медных 10-гигабитовых решений обеспечат еще большее снижение стоимости указанных соединений. Фактически речь идет о расширении и удешевлении полосы пропускания вычислительных сетей, основанных на более дешевых (по сравнению с оптическими) медных кабелях из витых пар, которые к тому же проще устанавливать и обслуживать. На сегодняшний день большинство проложенных каналов 10-*Gigabit Ethernet* работают как линии связи между коммутаторами. Увеличивая число медных портов на линейных платах, производители коммутаторов могут повысить плотность портов на 50% и снизить их стоимость, что также будет стимулировать заказчиков использовать

медные СКС, которые дешевле волоконно-оптических решений и твинаксиальных кабелей, предназначенных для работы с интерфейсами 10*GBase-CX*4.

Очевидно, что стандарт 10*GBase-T* пополнит семейство спецификаций физического уровня *Ethernet* (на базе медного кабеля из витых пар) и обеспечит очередное десятикратное (по сравнению с технологией 1000*Base-T*) повышение скорости передачи данных по технологии *Ethernet.* Это позволит сетевым администраторам повысить пропускную способность своих сетей до 10 Гбит/с, сохранив инвестиции в установленную медную кабельную инфраструктуру. Заметим, что в настоящее время на долю медной проводки приходится более 90 % всех инсталляций. Применение стандарта 10*GBase-T* должно снизить стоимость соединений 10 *Gigabit Ethernet* на 50 − 80 % за счет упрощения конструкции кабельной системы и упрощение ее инсталляции. Технология 10*GBase-T* позволит ускорить распространение 10 Гбит/с продуктов для кластерных многопроцессорных вычислительных систем.

Стандарт 10*GBase-T* призван стимулировать использование медных кабелей в локальных сетях для конструирования модульных многопроцессорных вычислительных систем.

Сетевой адаптер. В настоящее время выбор сетевых адаптеров для технологии 10 *Gigabit Ethernet* не очень велик. Заслуживают внимание разработки корпорации *Intel*, которая выпускает новые серверные сетевые адаптеры, работающие со скоростью 10 Гбит/с. Такие разработки впервые преодолели экономические и технические барьеры, делавшие в прошлом нерентабельным использование в центрах обработки данных каналов со скоростью 10 Гбит/с. После детального анализа такого рода адаптеров предпочтение было отдано двухпортовому адаптеру *Intel X*520-*T*2 [31]. Данное устройство поддерживает стандарт 10*Base-T*, т.е. обеспечивает реализацию интерфейса 10 *Gigabit Ethernet*. Сетевой адаптер *Intel X*520-*T*2 обеспечивает поддержку спецификации *SR-IOV* для расширенной сетевой виртуализации. При этом адаптер *Intel X*520-*T*2 является достаточно энергоэффективным устройством, что дает ему возможность поддерживать сразу два порта 10

Gigabit Ethernet. Преимущество такого подхода заключается в реализации методики дублирования, т.е. в случае сбоя в работе одного из портов второй продолжает функционировать. Это позволяет повысить надежность канала передачи информации. Кроме того, адаптер *Intel X*520-*T*2 можно объединить в один виртуальный канал 20*GbE* с соответствующим удвоением ширины пропускания. Такое свойство рассматриваемого адаптера имеет чрезвычайно важное значение для конструирования многоканального режима работы многопроцессорных вычислительных систем. Что касается его конструктивных особенностей, то необходимо отметить, что адаптер *Intel Ethernet Server Adapter X*520-*T*2 построен на базе 10*GbE* контроллера *Intel* 82599 и поддерживает сетевое соединение на расстоянии до 100 метров.

Коммутаторы. Важнейшим элементом сетевого интерфейса кластерной системы являются устройства коммутации сетевых каналов. На сегодняшний день одним из лидеров в области разработок коммутирующих устройств для технологии 10 *Gigabit Ethernet* является фирма *Arista Networks*. В продуктовой линейке фирмы *Arista Networks* имеются коммутаторы 2 и 3-его уровня серии 7100*T* с 24 и 48 портами [32]. Для рассматриваемой кластерной многопроцессорной системы предлагается применять двадцатичетырех-портовый коммутатор 7124*T*. Каждая модель этой серии сетевых коммутаторов имеет порты со стандартными *RJ*-45 гнездами с поддержкой автоопределения (*auto-negotiating)* стандартов 1*GbE* и 10*GBASE-T* и дополнительно *uplink* порт *SFP*+. Фирма позиционирует свои коммутаторы для высокоскоростного доступа к серверам. Порты 10*GbE* имеют обратную совместимость со стандартными портами *Gigabit Ethernet*.

Заметим, что коммутаторы высокой плотности серии 7100 с портами 10*GBASE-T* являются самыми подходящими для конструирования многопроцессорных кластерных вычислительных систем, поскольку при переходе от 1 *Gigabit Ethernet* до 10 *Gigabit Ethernet* они не требуют новой кабельной системы. При этом коммутаторы *Arista* 7100*T* поддерживают работу

10*GBASE-T* на кабельных линиях витая пара категории 6*a* до 100 метров, а также поддерживают работу на кабеле витая пара категории 6 до 55 метров.

Используя коммутаторы *Arista* 7100*T* и сетевые адаптеры 10*GBASE-T* такой компании, как *Intel*, можно обеспечивать высокоскоростное подключение в многопроцессорных вычислительных системах с использованием кабеля витая пара за приемлемую цену. Это решение позволяет развивать современные технологии в области высокопроизводительных компьютерных вычислений (*High Performance Computing HPC*), объединяя вычислительные мощности в кластеры или в вычислительные сети (*grid-based computing*).

Технические характеристики сетевого оборудования конструируемой модульной многопроцессорной кластерной системы приведены в табл. 2.13.

Таблица 2.13. Технические характеристики сетевого оборудования кластерной системы для технологии 10Gb Ethernet

Сетевой кабель	Тип	КС класса E/Cat 6
	Пропускная способность	10 Гбит/с
	Стандарт	*TIA/EIA*-568-*B*
	Цена	$ 25
Сетевой адаптер	Тип	*Server Adapter X520-T2*
	Производитель	*Intel*
	Пропускная способность	10 Гбит/с
	Цена	$ 1000
Коммутатор	Тип	7120*T*-4*S*
	Производитель	*Arista*
	Пропускная способность	480 Гб/с
	Цена за порт	$ 500 *port*

2.3.2. Исследование основных сетевых характеристик многопроцессорной системы для технологии 10*Gb Ethernet*

В соответствии с выбранным типом оборудования (табл. 2.13.), на первом этапе были проведены вычислительные эксперименты с целью определения основных сетевых характеристик кластерной системы. Такие исследования направлены на проверку процедуры согласования устройств сетевого интерфейса кластерной системы.

В качестве основы для проведения соответствующих расчетов примем аналитические соотношения (1.1) – (1.4). При таких обстоятельствах выполним процедуру моделирования сетевых коэффициентов в зависимости от количества узлов кластерной системы.

Исходные данные для изучения области изменения коэффициентов сетевого интерфейса многопроцессорной системы перечислены в табл. 2.14.

Таблица 2.14. Исходные данные для расчета сетевых характеристик кластерной системы для технологии 10Gb Ethernet

V_p	10 Гбит/с
V_b	480 Гбит/с

На первом этапе исследований, в соответствии с выражением (1.7), определим равновесное число узлов кластерной системы.

Проведенные расчеты показали, что точка сетевого равновесия соответствует числу узлов кластерной системы, равному $N = 24$. Это означает, что рассматриваемая кластерная система будет предоставлять более широкие возможности для решения поставленной задачи: уменьшения времени расчетов путем увеличения числа узлов кластерной системы. С другой стороны, можно отметить правильность выбора сетевого оборудования, т.к. равновесное число узлов кластерной системы полностью согласуется с техническими возможностями выбранного коммутатора. На этом этапе исследований можно отметить, что для рассматриваемого режима работы вычислительной сети коммутатор подобран удачно.

Далее, для уточнения особенностей функционирования сетевого интерфейса кластерной системы, на основании соотношений (1.3) – (1.5), выполним процедуру моделирования основных его числовых характеристик.

Полученные результаты моделирования сведены в табл. 2.15, а их геометрическая интерпретация представлена на рис. 2.9.

Таблица 2.15. Результаты расчета основных сетевых коэффициентов кластерной системы для технологии 10Gb Ethernet

Колич, узлов, N	k_s	K_b	c_k
1,00	0,04	24,00	480,00
2,00	0,08	12,00	240,00
3,00	0,13	8,00	160,00
4,00	0,17	6,00	120,00
5,00	0,21	4,80	96,00
6,00	0,25	4,00	80,00
7,00	0,29	3,43	68,57
8,00	0,33	3,00	60,00
9,00	0,38	2,67	53,33
10,00	0,42	2,40	48,00
11,00	0,46	2,18	43,64
12,00	0,50	2,00	40,00
13,00	0,54	1,85	36,92
14,00	0,58	1,71	34,29
15,00	0,63	1,60	32,00
16,00	0,67	1,50	30,00
17,00	0,71	1,41	28,24
18,00	0,75	1,33	26,67
19,00	0,79	1,26	25,26
20,00	0,83	1,20	24,00
21,00	0,88	1,14	22,86
22,00	0,92	1,09	21,82
23,00	0,96	1,04	20,87
24,00	1,00	1,00	20,00
25,00	1,04	0,96	19,20
26,00	1,08	0,92	18,46
27,00	1,13	0,89	17,78
28,00	1,17	0,86	17,14
29,00	1,21	0,83	16,55
30,00	1,25	0,80	16,00
31,00	1,29	0,77	15,48
32,00	1,33	0,75	15,00
33,00	1,38	0,73	14,55
34,00	1,42	0,71	14,12
35,00	1,46	0,69	13,71

Результаты проведенного моделирования основных сетевых коэффициентов кластерной системы для технологии 10 *Gigabit Ethernet* показывают, что при выбранном сетевом оборудовании существенно расширяется область дефицита сетевого интерфейса.

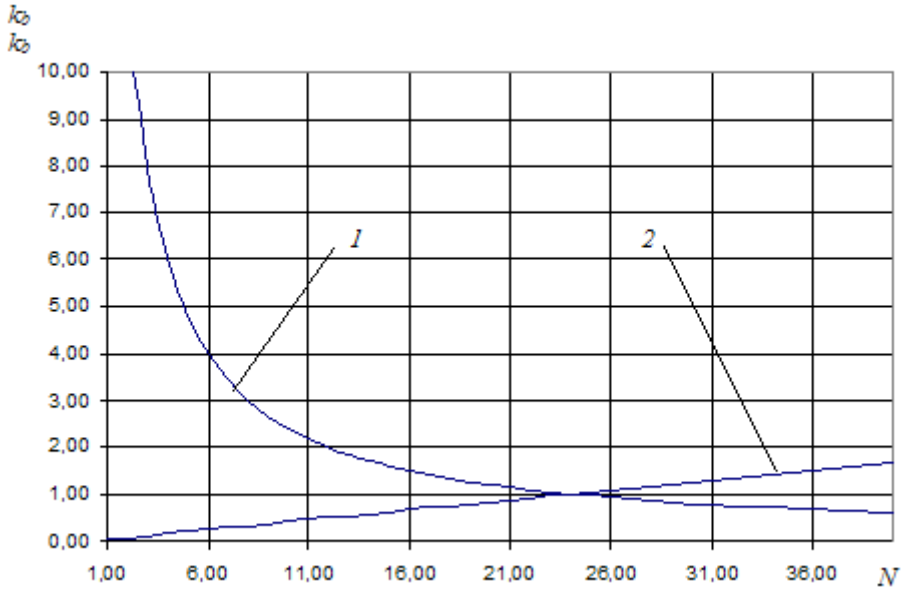

Рис. 2.9. Зависимости основных сетевых коэффициентов кластерной системы от количества узлов для технологии 10Gb Ethernet

В табл. 2.15 приведен расчет полосы пропускания коммутатора (c_k). Заметим, что для N = 24 значение полосы пропускания на каждый выходной порт кластерной системы будет соответствовать 20 Гбит/с, а это полностью согласуется как с дуплексным режимом обмена данных в многопроцессорной системе, так и с возможностями самой коммутационной матрицы. При таких обстоятельствах сетевой интерфейс многопроцессорной системы будет функционировать в режиме дефицита и в условиях максимально допустимой загрузки каналов коммутации коммутатора. Таким образом, можно отметить, что технические возможности предложенного коммутатора будут соответствовать протокольным возможностям сети. Следовательно, можно отметить, что сетевое оборудование выбрано удачно.

Итак, проведенный предварительный анализ результатов моделирования основных сетевых коэффициентов кластерной системы создал предпосылки для расчета основных характеристик ее эффективности.

2.3.3. Исследование оценок эффективности кластерной системы

Для анализа характеристик эффективности кластерной системы на первом этапе выполним расчет коэффициента пропускной способности кластерной системы (k_k). Расчет выполнялся на основании соотношений (1.4) и (1.8), а исходные данные соответствовали значениям, принятым в табл. 2.14.

График зависимости коэффициента пропускной способности кластерной системы от числа узлов представлен на рис. 2.10.

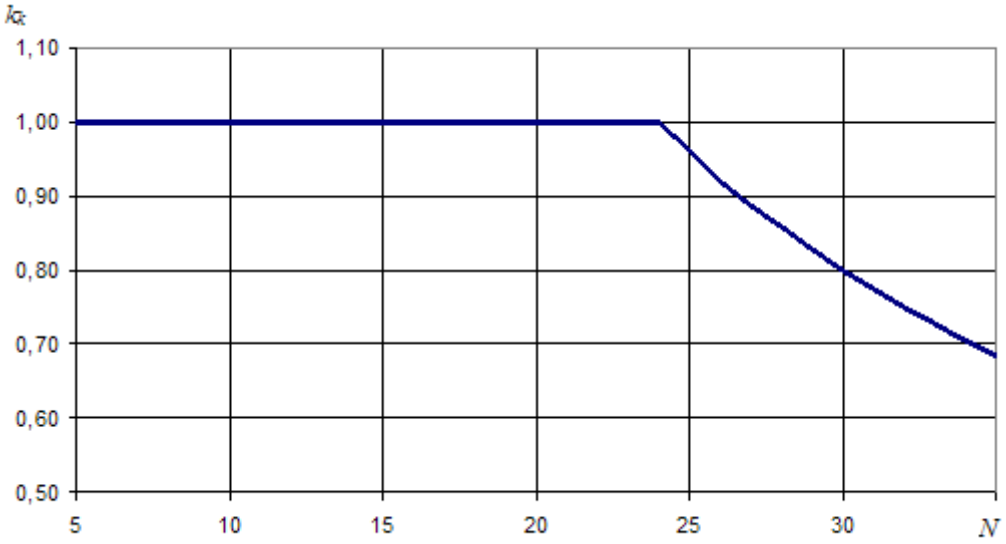

Рис. 2.10. График зависимости коэффициента пропускной способности кластерной системы от числа узлов для технологии 10Gb Ethernet

Анализ такой графической зависимости показывает, что здесь существенно расширяется область дефицита сетевого интерфейса (число узлов кластерной системы $N \le 24$). При таких условиях функционирования сети коммутационная матрица будет работать с наибольшей скоростью, т.е. передача данных в сети будет осуществляться в режиме сквозной коммутации (*cut-through*).

При числе узлов кластерной системы $N > 24$ коммутатор переходит в режим коммутации с буферизацией, что приводит к потере его

производительности, это обстоятельство и отражено убывающей линией на рис. 2.10.

Далее выполним процедуру исследования основных оценок эффективности кластерной системы. Исходные данные для исследования оценок эффективности кластерной системы представлены в табл. 2.16.

Таблица 2.16. Исходные данные для расчета характеристик эффективности многопроцессорной системы для технологии 10Gb Ethernet

V_p	10 Гбит/с
T_{it}	100 с
R	8 Гбит
m	2
d	2

На первом этапе для изучаемой кластерной многопроцессорной системы в условиях рассматриваемого эксперимента оценим количество узлов кластерной системы, при котором задача будет решаться наиболее эффективно. При этом воспользуемся аналитическими соотношениями вида (1.16) – (1.21).

Для работы кластера в режиме дефицита сетевого интерфейса воспользуемся уравнением (1.20). Решением такого уравнения будут два корня, при этом один из них положительный, а другой – отрицательный. Исходя из поставленных физических условий задачи, принимается положительный корень, значение которого равно девятнадцати, т.е. $N = 19$.

Для исследования работы кластера в режиме профицита сетевого интерфейса воспользуемся уравнением (1.21), которое будет иметь кубический вид. Решением такого уравнения будут два мнимых корня и один действительный. Действительный корень соответствует $N = 30$. Однако анализ такого корня показывает, что он не удовлетворяет условию функционирования кластерной системы в режиме профицита сетевого интерфейса (определение 1.6).

Отмеченные обстоятельства показывают, что в рамках рассматриваемой задачи оптимальное число узлов кластерной системы, при котором достигается

максимальная эффективность распараллеливания, будет соответствовать $N = 19$.

На втором этапе исследований выполним моделирование основных оценок эффективности кластерной системы. Данный этап реализован в соответствии с аналитическими соотношениями, выведенными в работе [10].

Полученные результаты моделирования сведены в табл. 2.17.

Таблица 2.17. Результаты расчета основных характеристик эффективности многопроцессорной системы для технологии 10Gb Ethernet

Колич. узлов, N	T_n	T_{ex}	*T*	*USK*	*EF*
1	100,00	0,00	100,00	1,00	1,00
2	50,00	0,28	50,28	1,99	0,99
3	33,33	0,57	33,90	2,95	0,98
4	25,00	0,85	25,85	3,87	0,97
5	20,00	1,13	21,13	4,73	0,95
6	16,67	1,41	18,08	5,53	0,92
7	14,29	1,70	15,98	6,26	0,89
8	12,50	1,98	14,48	6,91	0,86
9	11,11	2,26	13,37	7,48	0,83
10	10,00	2,55	12,55	7,97	0,80
11	9,09	2,83	11,92	8,39	0,76
12	8,33	3,11	11,44	8,74	0,73
13	7,69	3,39	11,09	9,02	0,69
14	7,14	3,68	10,82	9,24	0,66
15	6,67	3,96	10,63	9,41	0,63
16	6,25	4,24	10,49	9,53	0,60
17	5,88	4,53	10,41	9,61	0,57
18	5,56	4,81	10,36	9,65	0,54
19	5,26	5,09	10,35	9,66	0,51
20	5,00	5,37	10,37	9,64	0,48
21	4,76	5,66	10,42	9,60	0,46
22	4,55	5,94	10,49	9,54	0,43
23	4,35	6,22	10,57	9,46	0,41
24	4,17	6,51	10,67	9,37	0,39

Результаты моделирования представлены также в виде графических зависимостей (рис. 2.11, 2.12).

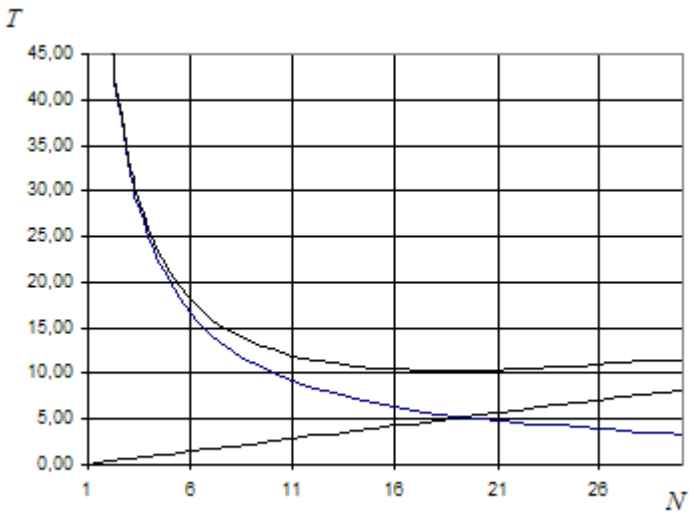

Рис. 2.11. Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы для технологии 10Gb Ethernet

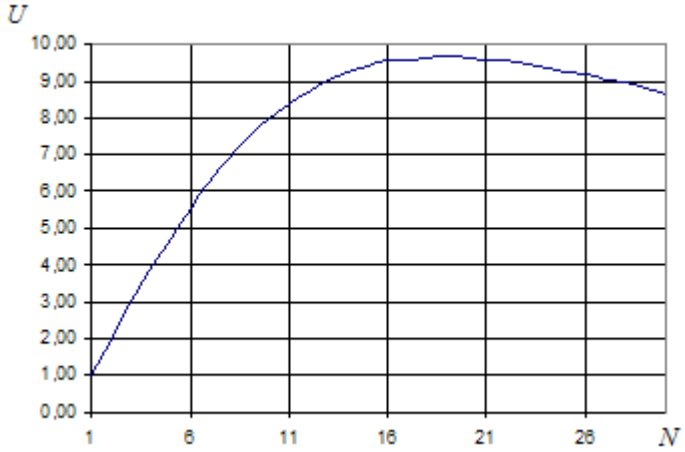

Рис. 2.12. Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы для технологии 10Gb Ethernet

Проведенный анализ полученных результатов моделирования показал следующее. Оптимальное число узлов кластерной системы, при котором будет достигаться максимальная эффективность распараллеливания, будет соответствовать N =19. Можно отметить, что ориентировочная цена такого сетевого оборудования будет составлять около 42000 у.е.

Наибольшая величина ускорения вычислений при предложенном сетевом интерфейсе соответствует значению, равному 9,66. Время счета задачи уменьшается со 100 с до 10,35 с.

2.3.4. Исследование загрузки вычислительной сети кластерной системы

Для анализа проверки правильности подобранного сетевого оборудования рассмотрим характеристику коэффициента использования сети многопроцессорной кластерной системы. С этой целью воспользуемся аналитическим выражением для расчета коэффициента использования сети, выведенного через параметры кластерной системы (1.23).

Результаты расчета коэффициента использования сети многопроцессорной системы приведены в табл. 2.18.

Таблица 2.18. Результаты расчета коэффициента использования сети кластера для технологии 10Gb Ethernet

Колич. узлов	*КЗС*
1	0,00
2	0,01
3	0,02
4	0,03
5	0,05
6	0,08
7	0,11
8	0,14
9	0,17
10	0,20
11	0,24
12	0,27
13	0,31
14	0,34
15	0,37
16	0,40
17	0,43
18	0,46
19	0,49
20	0,52
21	0,54
22	0,57
23	0,59
24	0,61

Полученные результаты позволяют сделать вывод, что для оптимального режима сетевого интерфейса необходимо использовать в многопроцессорной системе не более двадцати лезвий ($N \leq 20$). В данном случае показано, что

наилучшими оценки эффективности многопроцессорной системы будут, когда число лезвий многопроцессорной системы равно двенадцати ($N = 19$). Таким образом, можно отметить, что оборудование сетевого интерфейса кластерной системы подобрано удачно.

2.3.5. Многоканальный режим функционирования сетевого интерфейса

Для увеличения эффективной пропускной способности вычислительной сети кластерной системы рекомендуется использовать так называемое "связывание каналов" или режим агрегации каналов сетевого интерфейса. Такой способ объединения узлов кластерной системы в сеть отличается тем, что каждый узел подсоединяется к коммутатору более чем одним каналом. При этом главное преимущество режима агрегации каналов состоит в том, что радикально повышается скорость обмена данными в вычислительной сети кластерной системы. Заметим, что не каждый сетевой интерфейс можно связать в режим многоканального функционирования. Для этого комплектующие сетевого вычислительного канала должны удовлетворять ряду требований, которые детально изложены в разделе 1.7 настоящей монографии. Технология 10*GE* удовлетворяет таким требованиям. В этой связи интерес представляет многоканальный режим функционирования новой сетевой технологии.

2.3.5.1. Особенности подбора элементов сетевого интерфейса кластерной многопроцессорной системы для режима агрегации каналов сетевого интерфейса

Основная особенность технологии 10*GE* состоит в том, что для реализации многоканального интерфейса можно использовать те же комплектующие, что и для одного канала. Детальный анализ такого оборудования представлен в разделе 2.3.1 данной монографии. В то же время, особенности организации сетевого интерфейса для формирования режима агрегации каналов требуют наличия на каждом вычислительном узле двух

двухпортовых однотипных сетевых карт. Для рассматриваемой кластерной многопроцессорной системы предлагается применять 48 портовый коммутатор 7140*T*-8*S*. Более детально тип сетевого оборудования и его ориентировочная цена на сегодняшний день приведены в табл. 2.19.

Таблица 2.19. Технические характеристики сетевого оборудования кластерной системы в режиме агрегации каналов для технологии 10GE

Сетевой кабель	Тип	КС класса E/Cat 6
	Пропускная способность	10 Гбит/с
	Стандарт	*TIA/EIA*-568-*B*
	Цена	$ 25
Сетевой адаптер	Тип	*Server Adapter X520-T2*
	Производитель	*Intel*
	Пропускная способность	10 Гбит/с
	Цена	$ 1000
Коммутатор	Тип	7140*T*-8*S*
	Производитель	*Arista*
	Пропускная способность	800 Гб/с
	Цена за порт	$ 500 *port*

2.3.5.2. Исследование основных сетевых характеристик для режима агрегации каналов сетевого интерфейса

С учетом применения технологии агрегации каналов сетевого интерфейса определим основные сетевые характеристики кластерной системы, а именно, коэффициент пропускной способности сети кластера (k_s) и коэффициент пропускной способности коммутатора (k_b). Для этой цели воспользуемся соотношениями (1.24) и (1.25). Кроме того, для анализа согласования выбранной коммутационной шины с возможностями коммутатора, в соответствии с соотношением (1.26), определим коэффициент полосы пропускания коммутатора (c_k).

Исходные данные для изучения рассматриваемого режима работы сетевого интерфейса многопроцессорной системы перечислены в табл. 2.20.

Таблица 2.20. Основные сетевые характеристики кластерной системы для использования режима агрегации каналов сетевого интерфейса технологии 10GE

V_p	10 Гбит/с
V_b	800 Гбит/с
k	2
k_m	2

На первом этапе исследований, в соответствии с выражением (1.28), определим равновесное число узлов кластерной системы. С учетом заявленных возможностей сетевого интерфейса (табл. 2.20) равновесное число узлов кластерной системы соответствует $N = 40$.

Далее, для уточнения особенностей функционирования сетевого интерфейса кластерной системы, выполним процедуру моделирования основных его числовых характеристик. Полученные результаты моделирования сведены в табл. 2.21.

Результаты расчета основных сетевых коэффициентов кластерной системы с использованием высокопроизводительного сетевого интерфейса на рис. 2.13 представлены и в виде графических зависимостей.

Таблица 2.21. Результаты расчета основных сетевых коэффициентов кластера с использованием режима агрегации каналов сетевого интерфейса технологии 10GE

Колич, узлов N	k_s	k_b	c_k
1,00	0,03	40,00	1600,00
2,00	0,05	20,00	800,00
3,00	0,08	13,33	533,33
4,00	0,10	10,00	400,00
5,00	0,13	8,00	320,00
6,00	0,15	6,67	266,67
7,00	0,18	5,71	228,57
8,00	0,20	5,00	200,00
9,00	0,23	4,44	177,78
10,00	0,25	4,00	160,00
11,00	0,28	3,64	145,45
12,00	0,30	3,33	133,33
13,00	0,33	3,08	123,08
14,00	0,35	2,86	114,29
15,00	0,38	2,67	106,67
16,00	0,40	2,50	100,00
17,00	0,43	2,35	94,12
18,00	0,45	2,22	88,89
19,00	0,48	2,11	84,21
20,00	0,50	2,00	80,00
21,00	0,53	1,90	76,19
22,00	0,55	1,82	72,73
23,00	0,58	1,74	69,57
24,00	0,60	1,67	66,67
25,00	0,63	1,60	64,00
26,00	0,65	1,54	61,54
27,00	0,68	1,48	59,26
28,00	0,70	1,43	57,14
29,00	0,73	1,38	55,17
30,00	0,75	1,33	53,33
31,00	0,78	1,29	51,61
32,00	0,80	1,25	50,00
33,00	0,83	1,21	48,48
34,00	0,85	1,18	47,06
35,00	0,88	1,14	45,71
36,00	0,90	1,11	44,44
37,00	0,93	1,08	43,24
38,00	0,95	1,05	42,11
39,00	0,98	1,03	41,03
40,00	1,00	1,00	40,00
41,00	1,03	0,98	39,02
42,00	1,05	0,95	38,10
43,00	1,08	0,93	37,21
44,00	1,10	0,91	36,36

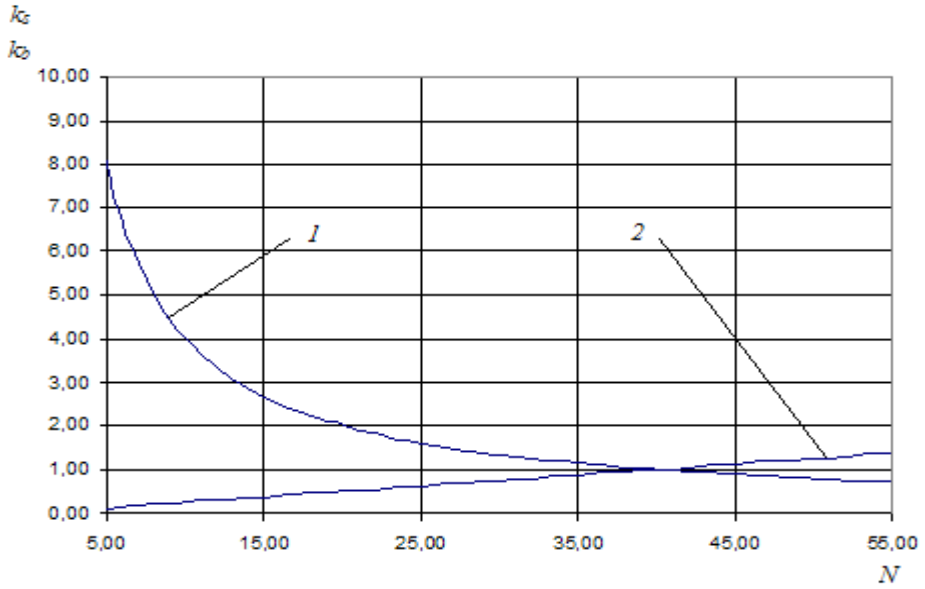

Рис. 2.13. Зависимости основных сетевых коэффициентов кластерной системы от количества узлов с использованием режима агрегации каналов сетевого интерфейса технологии 10GE

Проведем анализ полученных результатов. Тогда точка сетевого равновесия будет соответствовать $N = 40$ (рис. 2.13). Очевидно, что представленный режим работы, при равных прочих условиях, за счет изменения коммутационной матрицы и архитектуры сетевого интерфейса многопроцессорной системы позволил расширить не только полосу пропускания коммутационной шины и равновесное число узлов кластера. *Последнее обстоятельство означает, что сформированный режим работы сетевого интерфейса кластерной системы будет предоставлять более широкие возможности для реализации процесса обмена данными между вычислительными узлами многопроцессорной системы, существенно улучшая характеристики эффективности, быстродействия и надежности функционирования системы.*

В соответствии со сформированным режимом функционирования сетевого интерфейса возникают предпосылки для переоценки характеристик ускорения и эффективности кластерной системы.

2.3.5.3. Исследование оценок эффективности кластерной системы

В соответствии со сформированным режимом функционирования сетевого интерфейса проведем этап моделирования основных оценок эффективности кластерной системы.

Для всестороннего освещения процессов, протекающих в многопроцессорной вычислительной системе, на первом этапе исследований проведем анализ коэффициента пропускной способности кластерной системы (k_k). Расчет выполнялся на основании соотношений (1.8) и (1.25), а исходные данные соответствовали значениям, принятым в табл. 2.20.

График зависимости коэффициента пропускной способности кластерной системы от числа узлов представлен на рис. 2.14.

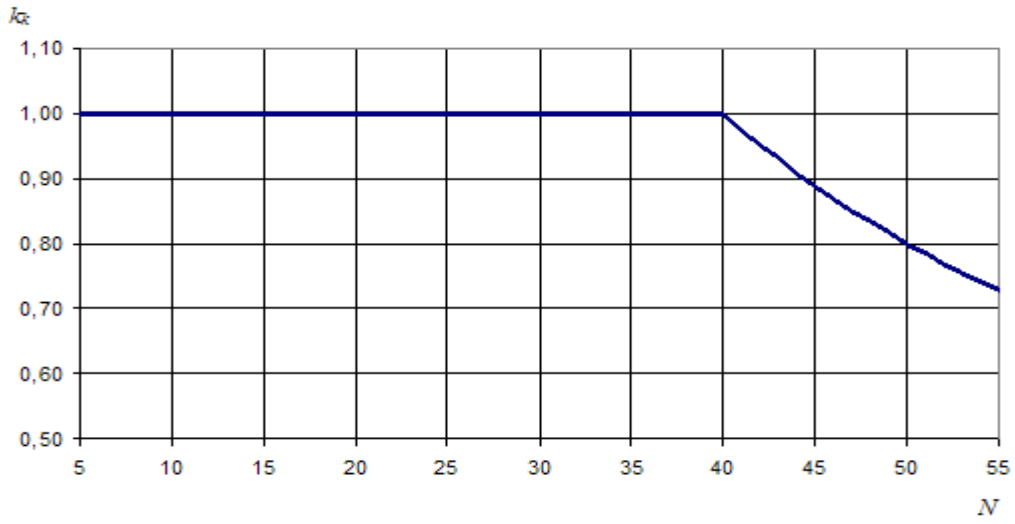

Рис. 2.14. График зависимости коэффициента пропускной способности кластерной системы от числа узлов режима агрегации каналов сетевого интерфейса технологии 10GE

Анализ такой графической зависимости показывает, что для режима дефицита сетевого интерфейса (когда число узлов кластерной системы $N \le 40$) коэффициент пропускной способности кластерной системы будет определяться характеристиками сетевого интерфейса. При таких условиях функционирования сетевого интерфейса коммутационная матрица будет передавать данные с наиболее возможной скоростью.

В режиме же профицита сетевого интерфейса (когда число узлов кластерной системы $N > 40$) такой параметр будет определяться характеристиками коммутатора, когда сумма входящих пакетов превышает сумму выходящих. Здесь коммутатор переходит в режим коммутации с задержкой фреймов в буфере, что приводит к потере его производительности, это обстоятельство и отражено убывающей линией на рис. 2.14.

Заметим, что элементы сетевого интерфейса кластерной системы выбирались с учетом высокопроизводительного режима технологии $10GE$. Такой подход позволяет существенно улучшить оценки эффективности модульной многопроцессорной кластерной системы. В таком случае особый интерес будут представлять оценки эффективности многопроцессорной вычислительной системы. Исходные данные для исследования оценок эффективности такой кластерной системы представлены в табл. 2.22.

Таблица 2.22. Исходные данные для расчета характеристик эффективности многопроцессорной системы в режиме агрегации каналов

V_p	10 Гбит/с
T_{it}	100 с
R	8 Гбит
m	2
d	2
k	2
k_m	2

Здесь k – количество симметричных вычислительных сетей, которые работают одновременно за счет реализации технологии агрегации каналов, k_m – количество коммутационных матриц в сети обмена данных.

На первом этапе для изучаемой кластерной многопроцессорной системы в условиях рассматриваемого эксперимента оценим количество узлов кластерной системы, при котором задача будет решаться наиболее эффективно. При этом воспользуемся аналитическими соотношениями вида (1.30) – (1.37).

Для работы кластера в режиме дефицита сетевого интерфейса воспользуемся уравнением (1.36). Решением такого уравнения будут два корня, при этом один из них положительный, а другой – отрицательный. Исходя из

поставленных физических условий задачи, принимается положительный корень, значение которого равно двенадцати, т.е. $N = 27$. Заметим, что такое решение удовлетворяет неравенству из определения 1.5, которое устанавливает условия функционирования кластерной системы в режиме дефицита сетевого интерфейса.

Для исследования работы кластера в режиме профицита сетевого интерфейса воспользуемся уравнением (1.37), которое будет иметь кубический вид. Решением такого уравнения будут два мнимых корня и один действительный. Действительный корень соответствует $N = 38$. Однако анализ такого корня показывает, что он не удовлетворяет условию функционирования кластерной системы в режиме профицита сетевого интерфейса (определение 1.6).

Отмеченные обстоятельства показывают, что в рамках рассматриваемой задачи оптимальное число узлов кластерной системы, при котором будет достигаться максимальная эффективность распараллеливания, будет соответствовать $N = 27$.

На втором этапе исследований выполним моделирование основных оценок эффективности кластерной системы. Данный этап реализован в соответствии с аналитическими соотношениями, выведенными в работе [10].

Полученные результаты моделирования сведены в табл. 2.23.

Результаты моделирования представлены также в виде графических зависимостей (рис. 2.15, 2.16).

Таблица 2.23. Результаты расчета основных характеристик эффективности при реализации двухканального режима функционирования вычислительной сети кластера по технологии 10GE

Колич. узлов, N	T_n	T_{ex}	*T*	*USK*	*EF*
1	100,00	0,00	100,00	1,00	1,00
2	50,00	0,14	50,14	1,99	1,00
3	33,33	0,28	33,62	2,97	0,99
4	25,00	0,42	25,42	3,93	0,98
5	20,00	0,57	20,57	4,86	0,97
6	16,67	0,71	17,37	5,76	0,96
7	14,29	0,85	15,13	6,61	0,94
8	12,50	0,99	13,49	7,41	0,93
9	11,11	1,13	12,24	8,17	0,91
10	10,00	1,27	11,27	8,87	0,89
11	9,09	1,41	10,51	9,52	0,87
12	8,33	1,56	9,89	10,11	0,84
13	7,69	1,70	9,39	10,65	0,82
14	7,14	1,84	8,98	11,13	0,80
15	6,67	1,98	8,65	11,57	0,77
16	6,25	2,12	8,37	11,95	0,75
17	5,88	2,26	8,15	12,28	0,72
18	5,56	2,40	7,96	12,56	0,70
19	5,26	2,55	7,81	12,81	0,67
20	5,00	2,69	7,69	13,01	0,65
21	4,76	2,83	7,59	13,17	0,63
22	4,55	2,97	7,52	13,31	0,60
23	4,35	3,11	7,46	13,41	0,58
24	4,17	3,25	7,42	13,48	0,56
25	4,00	3,39	7,39	13,52	0,54
26	3,85	3,54	7,38	13,55	0,52
27	3,70	3,68	7,38	13,55	0,50
28	3,57	3,82	7,39	13,53	0,48
29	3,45	3,96	7,41	13,50	0,47
30	3,33	4,10	7,43	13,45	0,45
31	3,23	4,24	7,47	13,39	0,43
32	3,13	4,38	7,51	13,32	0,42
33	3,03	4,53	7,56	13,23	0,40
34	2,94	4,67	7,61	13,14	0,39
35	2,86	4,81	7,67	13,05	0,37
36	2,78	4,95	7,73	12,94	0,36
37	2,70	5,09	7,79	12,83	0,35
38	2,63	5,23	7,86	12,72	0,33
39	2,56	5,37	7,94	12,60	0,32
40	2,50	5,52	8,02	12,48	0,31

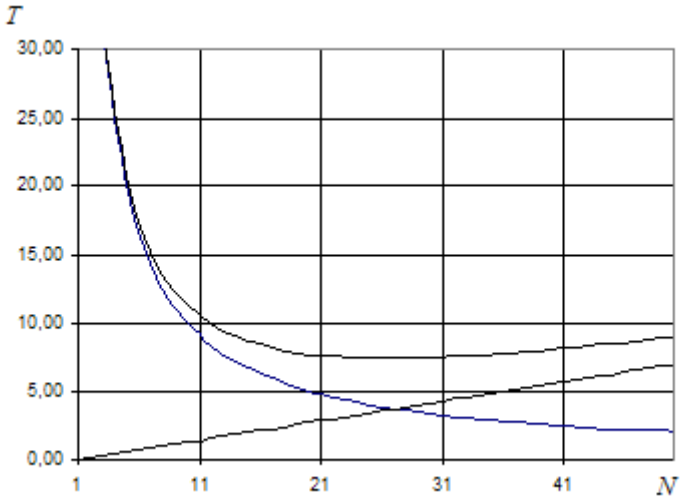

Рис. 2.15. Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы для технологии 10Gb Ethernet в режиме агрегации каналов

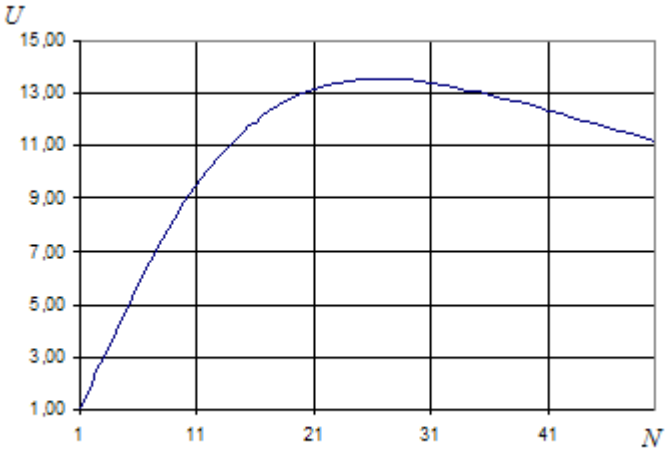

Рис. 2.16. Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы для технологии 10Gb Ethernet в режиме агрегации каналов

Проведенный анализ полученных результатов моделирования показал следующее. Оптимальное число узлов кластерной системы, при котором будет достигаться максимальная эффективность распараллеливания, будет соответствовать $N = 27$. Ориентировочная цена такого сетевого оборудования будет составлять около 115000 у.е.

Наибольшая величина ускорения вычислений при предложенном сетевом интерфейсе соответствует значению, равному 13,55. Время счета задачи уменьшается со 100 с до 7,38 с.

2.3.5.4. Исследование загрузки вычислительной сети кластерной системы

Для анализа проверки правильности подобранного сетевого оборудования рассмотрим характеристику коэффициента использования сети многопроцессорной кластерной системы. С этой целью воспользуемся аналитическим выражением для расчета коэффициента использования сети, выведенного через параметры кластерной системы (1.40).

Результаты расчета коэффициента использования сети многопроцессорной системы приведены в табл. 2.24.

Таблица 2.24. Результаты расчета коэффициента использования сети кластера для технологии 10Gb Ethernet в режиме агрегации каналов

Колич. узлов	*КЗС*
1	0,00
2	0,00
3	0,01
4	0,02
5	0,03
6	0,04
7	0,06
8	0,07
9	0,09
10	0,11
11	0,13
12	0,16
13	0,18
14	0,20
15	0,23
16	0,25
17	0,28
18	0,30
19	0,33
20	0,35
21	0,37
22	0,40

23	0,42
24	0,44
25	0,46
26	0,48
27	0,50
28	0,52
29	0,53
30	0,55

Полученные результаты позволяют сделать вывод, что, для оптимального режима сетевого интерфейса необходимо использовать в многопроцессорной системе не более двадцати семи лезвий ($N \leq 27$). В данном случае показано, что наилучшими оценки эффективности многопроцессорной системы будут, когда число лезвий многопроцессорной системы равно двадцати семи ($N = 27$). Таким образом, можно отметить, что оборудование сетевого интерфейса кластерной системы подобрано удачно.

Раздел 3.
АНАЛИЗ РАЗВИТИЯ И ПЕРСПЕКТИВ ПРИМЕНЕНИЯ СЕТЕВЫХ ИНТЕРФЕЙСОВ МНОГОПРОЦЕССОРНЫХ СИСТЕМ

Анализ исследования нынешнего состояния и перспектив применения современных коммуникационных технологий в многопроцессорных кластерных системах проводился на основании учета следующих характеристик: ускорение вычислений, время счета задачи и цена сетевого интерфейса. В табл. 3.1 приведены оценки влияния ценового фактора современных коммуникационных технологий на эффективность распараллеливания и время решения задачи.

Для более детального анализа полученных результатов приведем основные графические зависимости, иллюстрирующие оценки влияния ценового фактора современных коммуникационных технологий, эффективность распараллеливания и время решения задачи от количества узлов кластерной системы. Так, на рис. 3.1. приведен график зависимости цены сетевого интерфейса от числа узлов кластерной системы. Цифрой на соответствующих маркерах указана скорость сетевого интерфейса.

Таблица 3.1. Оценки влияния ценового фактора современных коммуникационных технологий на эффективность распараллеливания и время решения задачи

Тип техно-логии	*Колич. узлов, N*	*Скорость сети, Гб/с*	*Время счета, с.*	*Уско-рение*	*Цена сетевого интерфейса, у.е.*
GE	6	1	30,81	3,25	3000
GE+CB	9	2	22,4	4,60	5000
FC-4	12	4	16,11	6,21	34000
FC-8	17	8	11,54	8,67	50000
10*GbE*	19	10	10,35	9,66	42000
10*GbE+CB*	27	20	7,38	13,55	110000

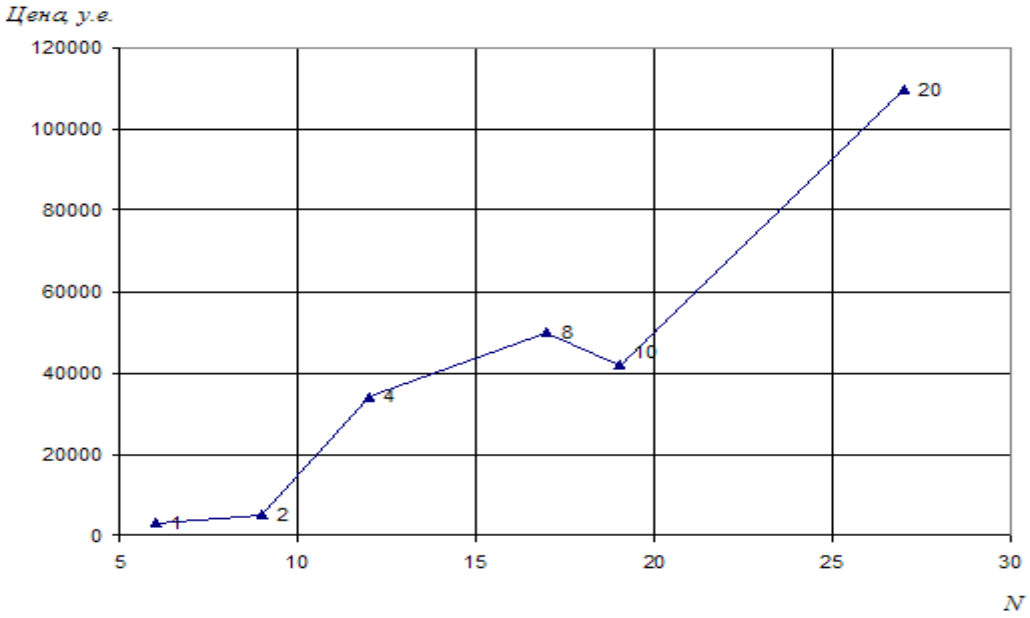

Рис. 3.1. График зависимости цены сетевого интерфейса от числа узлов кластерной системы

Такой график иллюстрирует перспективность применения сетевых архитектур *Ethernet.* Еще совсем недавно казалось, что благодаря высокой скорости передачи данных, малой задержке и расширяемости сетевая технология *Fibre Channel* (*FC*) практически не имеет аналогов в своей области. Более того, в последние годы сетевая технология *FC* все чаще используется для конструирования высокопроизводительных вычислительных систем. Однако принятие стандарта 10*GBase-T* стимулирует использование медных кабелей при проектировании многопроцессорных вычислительных систем.

Заметим, что на сегодняшний день, на долю медной проводки приходится более 90% всех инсталляций. Применение стандарта 10*GBase-T* должно снизить стоимость соединений 10 *Gigabit Ethernet* на 50 − 80 % за счет упрощения конструкции кабельной системы и облегчения ее инсталляции. Сетевая технология 10*GBase-T* позволяет ускорить распространение 10*GE* для проектирования многопроцессорных кластерных систем.

Увеличивая число медных портов на линейных платах, производители коммутаторов могут повысить плотность портов на 50 % и снизить их стоимость, что также будет стимулировать заказчиков использовать медные

СКС, которые дешевле волоконно-оптических решений. Фактически речь идет о расширении и удешевлении полосы пропускания сетевого интерфейса кластерных вычислительных систем, основанных на более дешевых (по сравнению с оптическими) медных кабелях из витых пар, которые к тому же проще устанавливать и обслуживать.

Отмеченные замечания иллюстрирует графическая зависимость, представленная на рис. 3.1. Так, здесь показано, что на сегодняшний день сетевая технология 10*GE,* по сравнению с оптоволоконным каналом *Fibre Channel.* позволяет получить соответствующие преимущества по ценовому фактору и оценкам эффективности.

На рис. 3.2. приведен график зависимости времени счета задачи от числа узлов кластерной системы. Цифрой на соответствующих маркерах указана скорость сетевого интерфейса. График имеет выраженную гиперболическую зависимость и позволяет прогнозировать не только время счета задач в многопроцессорных кластерных системах, но и дальнейшее развитие сетевого интерфейса.

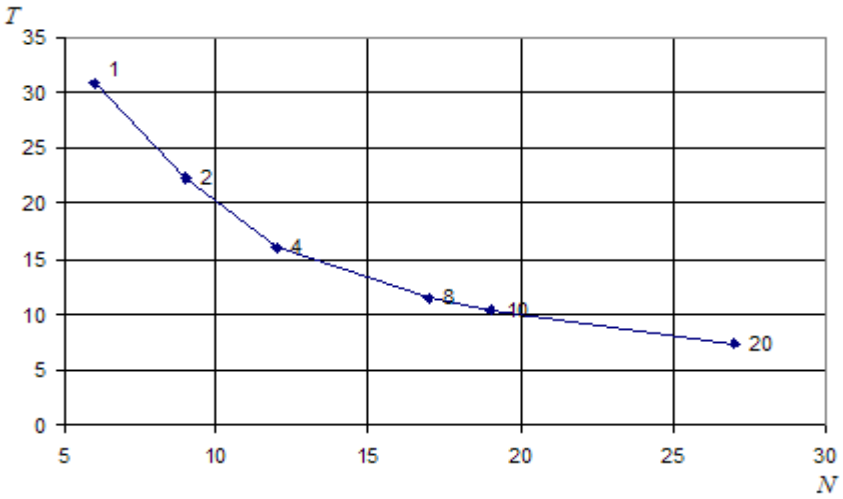

Рис. 3.2. График зависимости времени счета задачи от числа узлов кластерной системы

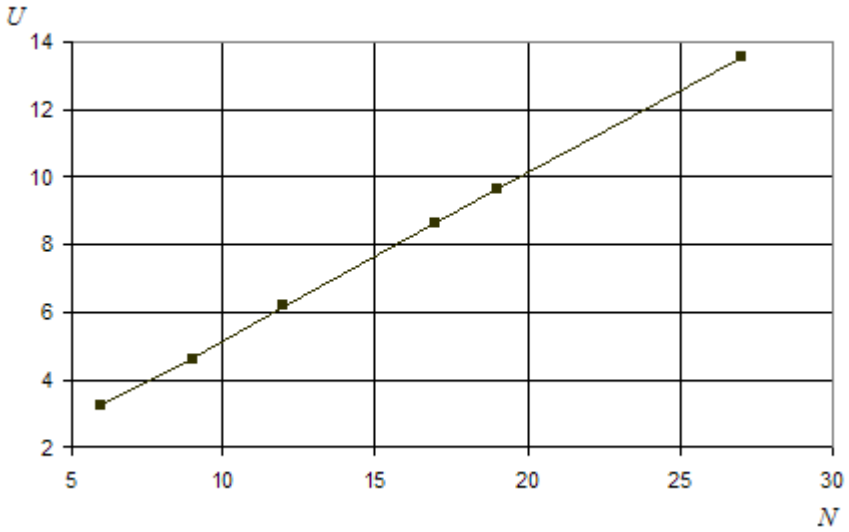

Рис. 3.3. График зависимости ускорения вычислений от числа узлов кластерной системы

Наконец, на рис. 3.3 приведен график зависимости ускорения вычислений от числа узлов кластерной системы. Маркерами на рисунке указана величина ускорения для соответствующих технологий. Особый интерес здесь представляет линейный характер такой зависимости. Такой график позволяет сделать соответствующий прогноз относительно оценок ускорения в многопроцессорных вычислительных системах.

Раздел 4.
ОСОБЕННОСТИ СОПРЯЖЕНИЯ МОДУЛЬНЫХ МНОГОПРОЦЕССОРНЫХ КЛАСТЕРНЫХ СИСТЕМ

В конфигурации рассматриваемой многопроцессорной кластерной системы [2] было избрано шесть лезвий и модульный принцип ее реализации. Это обеспечивает в случае необходимости расширения установку дополнительных модулей. При этом существует широкий класс сильносвязанных задач [10, 11], для решения которых при помощи модульной многопроцессорной системы в режиме ее наибольшей эффективности шести лезвий может оказаться недостаточно. В этой связи, данный раздел монографии направлен на решение проблемы сопряжения модульных многопроцессорных кластерных систем.

Итак, рассматривается задача сопряжения модулей многопроцессорных вычислительных систем. На рис. 4.1 представлена схема сопряжения интерфейсов двух модулей (Модуль 1 и узел расширения – Модуль 2) многопроцессорных вычислительных систем. Рассматриваемые модули однотипны. Так, в модуле 1 представлен мастер – узел (*MNode*001) и пять *slave* – узлов (*Node*001, *Node*002, *Node*003, *Node*004, *Node*005). Такой модуль также содержит управляемый коммутатор (УК1), образующий сеть управления загрузки и диагностики. Обмен данными между вычислительными узлами многопроцессорной системы вынесен в отдельную сеть с использованием механизма *channel bonding* и *VPN*, что позволило увеличить скорость обмена данными и снизить загрузку канала между узлами кластера. Агрегация каналов реализована при помощи коммутаторов УК2 и УК3 по топологии сети звезда.

Применение энергоэффективных вычислительных узлов модуля кластерной системы позволило запитать все *slave* – узлы от одного блока питания (ATX2), уменьшив тем самым потребление энергии и увеличив надежность вычислительной системы в целом. Мастер – узел запитывается от индивидуального блока питания (ATX1). Запуск и инициализация модуля многопроцессорной системы осуществляются при помощи панели управления

(П 01). В мастер – узле и *slave* – вычислительных узлах применяются одни и те же комплектующие (материнские платы, процессоры, внешние сетевые платы, мастер-узел оборудован дополнительно жестким диском, *CD-ROM*, дисководами).

Аналогичную конфигурацию имеет и модуль расширения (модуль 2).

Сетевые интерфейсы модуля 1 организованы по следующему принципу. Интегрированный сетевой интерфейс (*inc*001) узла *MNode*001 с функцией сетевой загрузки подсоединен входом/выходом к порту 06 управляемого коммутатора УК1. Интегрированный сетевой интерфейс *NNode*001_*inc*001 узла *NNode*001 подсоединен входом/выходом к порту 01 управляемого коммутатора УК1. Аналогично выполнено соединение остальных узлов кластерной системы с управляемым коммутатором УК1.

Архитектура вычислительной сети кластера для реализации режима *channel bonding* реализована следующим образом. *Slave* – узел *NNode*001 соединен входом/выходом двунаправленным внешним сетевым интерфейсом *N*001,1 *Gi*_001,1 с 01 портом управляемого коммутатора УК2. Кроме того, дополнительно такой узел соединен двунаправленным сетевым интерфейсом *N*001,2 *Gi*_001,2 с 01 портом управляемого коммутатора УК3.

Аналогично выполнено соединение остальных *slave*-узлов кластерной системы с управляемыми коммутаторами УК2 и УК3.

Режим сопряжения модуля 1 и модуля 2 реализован следующим образом. Сопряжение сетей управления выполнено на управляемых коммутаторах УК1 и УК4. При этом порт 08 УК1 соединен с портом 07 УК4. Режим коммутации управляемых коммутаторов УК1 и УК4 реализуется при помощи кросс-соединения указанных портов. Такой подход позволяет объединить сети управления модуля 1 и модуля 2 для обеспечения режима загрузки, диагностики и управления многопроцессорной кластерной системы в целом.

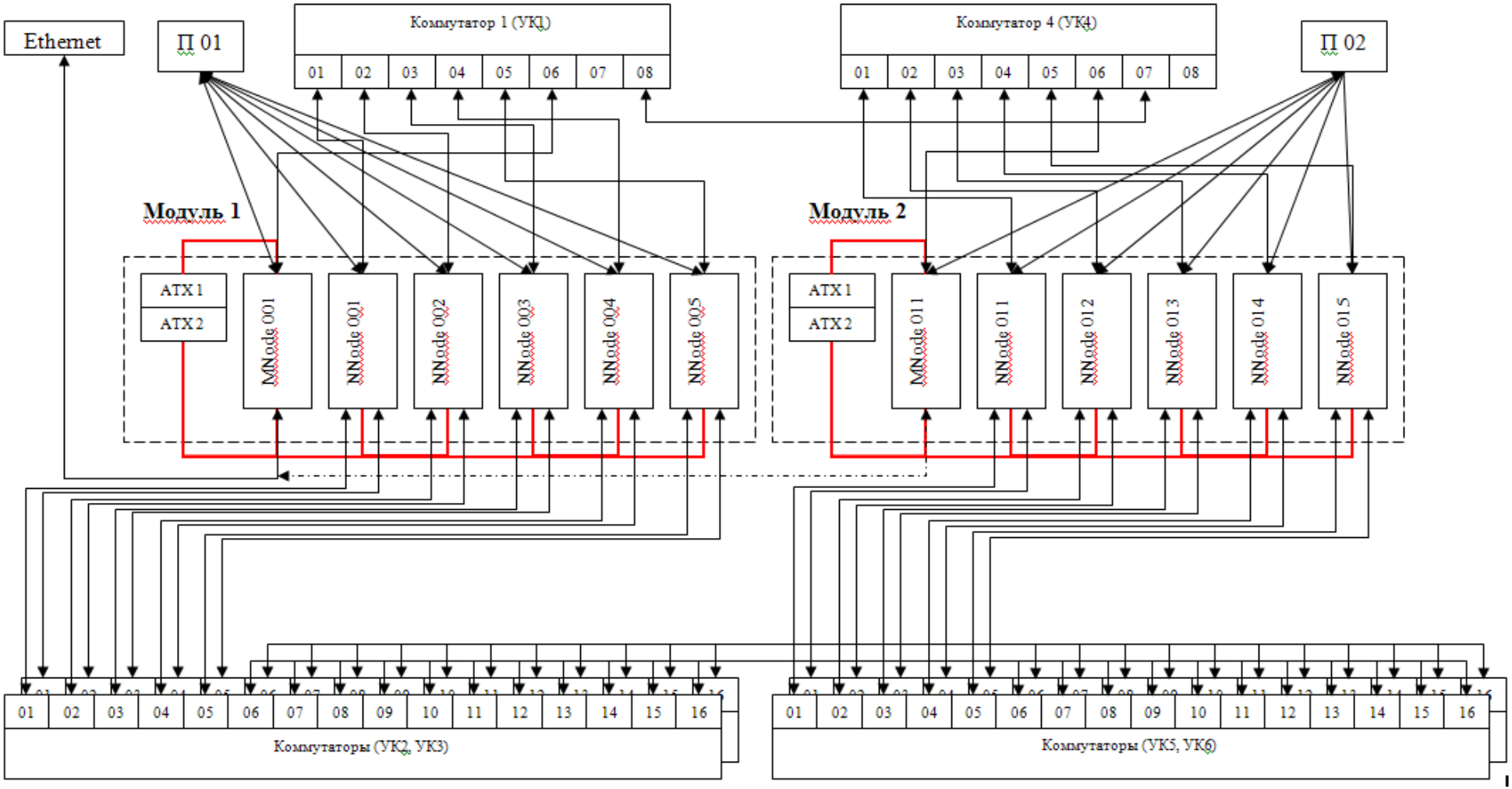

Рис. 4.1. Схема сопряжения интерфейсов модулей многопроцессорных вычислительных систем

Сопряжение сетевых интерфейсов в сети обмена данными модуля 1 и модуля 2 реализовано по следующей схеме. Управляемый коммутатор УК2 соединяется своими портами с управляемым коммутатором УК5, а управляемый коммутатор УК3 соединяется своими портами с управляемым коммутатором УК6. Коммутация портов указанных коммутаторов осуществляется следующим образом. Порт 06 управляемого коммутатора УК2 (УК3) соединен с портом 06 управляемого коммутатора УК5 (УК6), аналогично соединяются одноименные порты таких коммутаторов с порта 07 по порт 16. Такой подход позволяет совместить сети обмена данными модулей многопроцессорных систем без потери пропускной способности.

Работа вычислительного комплекса начинается с поступления внешнего сигнала *PUSK* панели управления П01 модуля 1 на мастер – узел *MNode*001. Загрузка ОС мастер – узла осуществляться с жесткого диска либо CD/DVD – устройства. Затем запускается конфигурационный скрипт, который настраивает работу *DHCP* сервера. Кроме того, здесь указывается количество вычислительных узлов системы, в случае необходимости разрешается доступ в Интернет или внешнюю сеть, указываются основные настройки и ее параметры.

Мастер – узел *MNode*001 по сети загрузки и управления (коммутаторы УК1 и УК4) осуществляет инициализацию и загрузку *slave* – узлов *NNode*001 – NNode005, NNode011 – NNode015 модулей вычислительной системы. Такой подход позволяет конфигурировать и перенастраивать кластерную систему в оптимальном режиме и в сжатые сроки. Принимая во внимание то обстоятельство, что ОС вычислительной системы загружается по сети, появилась необходимость запускать по пять вычислительных *slave* – узлов одновременно. Это потребовало реализации режима *Power on After Power Fail / Former-Sts*. По сети управления *slave* – узлы получают соответствующие настройки, загружаются и ожидают ответа от *DHCP* сервера. После соответствующей загрузки команд начинается непосредственно работа сопряженной многопроцессорной кластерной системы. Режим коммутации

управляемых коммутаторов УК2, УК3 и УК5, УК6 реализуется при помощи кросс-панели.

Для сильносвязанных задач в комплексной многопроцессорной системе реализован режим обмена данными по топологии кольцо. Схема формирования такого режима представлена на рис. 4.2.

Для исключения взаимного влияния при передаче/приеме данных между вычислительными узлами применяются *VPN* сети внутри управляемых коммутаторов УК2 и УК5. *Slave* – узел *NNode*001 соединен входом/выходом двунаправленным внешним сетевым интерфейсом *N*001,1 *Gi*_001,1 с 10 портом управляемого коммутатора УК5. Кроме того, такой узел дополнительно соединен двунаправленным сетевым интерфейсом *N*001,2 *Gi*_001,2 с 01 портом управляемого коммутатора УК3. *Slave* – узел *NNode*002 соединен входом/выходом двунаправленным внешним сетевым интерфейсом *N*002,1 *Gi*_002,1 с 02 портом управляемого коммутатора УК2. Кроме того, такой узел дополнительно соединен двунаправленным сетевым интерфейсом *N*002,2 *Gi*_002,2 с 03 портом управляемого коммутатора УК2. В соответствии с приведенным подходом соединяются и остальные вычислительные узлы многопроцессорной системы, образуя топологию кольцо.

В коммутаторах УК2 и УК5 создаются по пять виртуальных сетей: *vpn*01.2 между портами 01 и 02 коммутатора УК2, *vpn*02.3 между портами 03 и 04 коммутатора УК2, *vpn*03.4 между портами 05 и 06 коммутатора УК2, *vpn*04.5 между портами 07 и 08 коммутатора УК2, *vpn*05.6 между портами 09 и 10 коммутатора УК5, *vpn*11.2 между портами 01 и 02 коммутатора УК5, *vpn*12.3 между портами 03 и 04 коммутатора УК5, *vpn*13.4 между портами 05 и 06 коммутатора УК5, *vpn*14.5 между портами 07 и 08 коммутатора УК5, *vpn*15.6 между портами 09 и 10 коммутатора УК5.

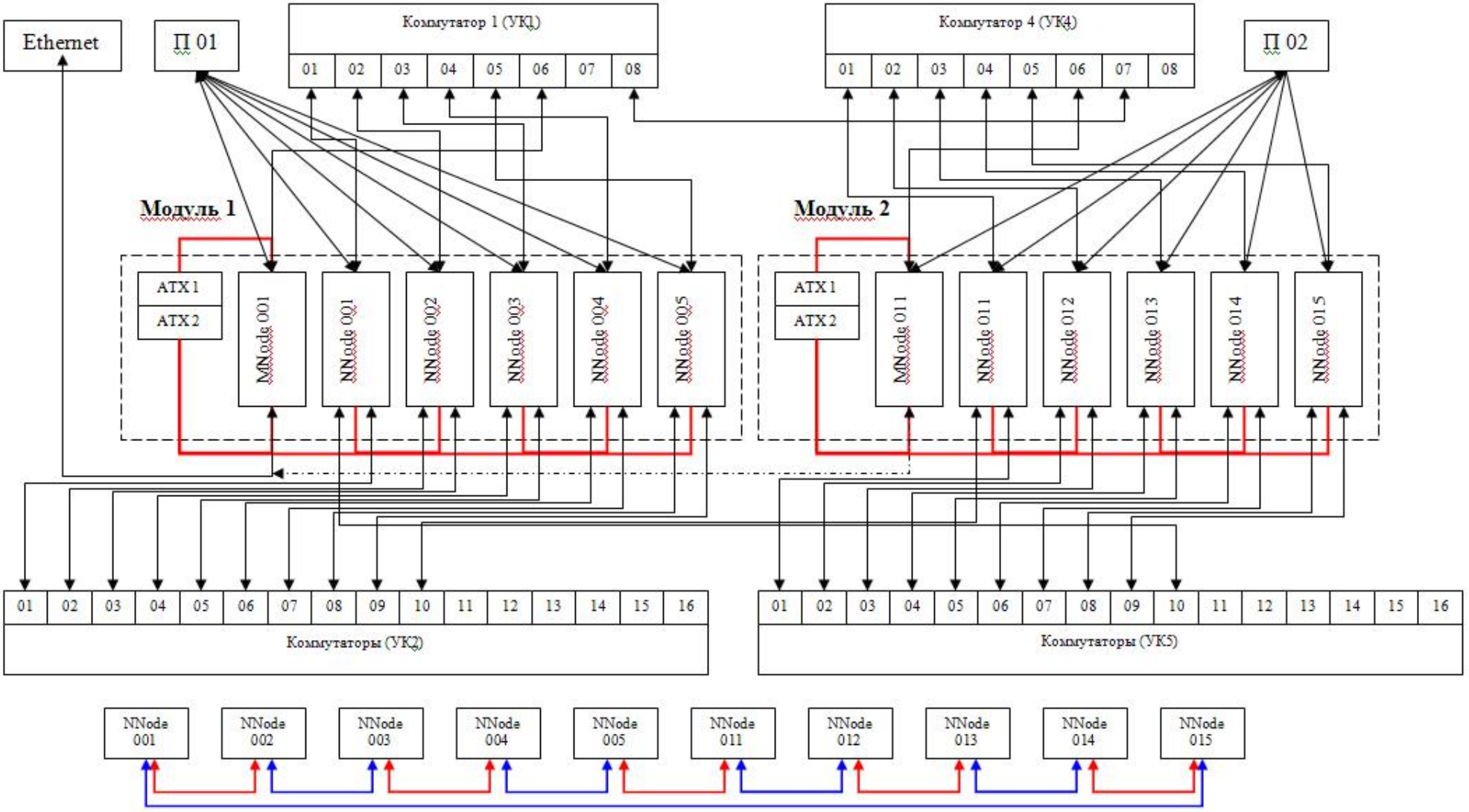

Рис. 4.2. Структура сети многопроцессорной системы при реализации граничного обмена данными по топологии кольцо

Для сконфигурированного вычислительного комплекса был проведен ряд вычислительных экспериментов. Эффективность предложенного подхода при проведении вычислительных работ подтверждается решением задач нестационарной теплопроводности, моделированием обратных задач исследования теплофизических свойств материалов, моделированием задач прогноза экологических систем, находящихся под воздействием естественных и антропогенных факторов, моделированием температурного режима длинномерного изделия при его термической обработке.

Выводы

В монографии показаны пути повышения эффективности многопроцессорной кластерной системы за счет реорганизации архитектуры ее сетевого интерфейса. Предложенный подход позволил не только повысить эффективность распараллеливания и надежность функционирования кластерной системы, но и существенно уменьшить время вычислений. Таких результатов удалось достичь за счет уменьшения времени граничного обмена данными между вычислительными узлами кластерной системы.

При этом:

1. Выявлены два основных режима работы сетевого интерфейса в модульной многопроцессорной кластерной системе. Показаны условия формирования равновесного числа вычислительных узлов многопроцессорной кластерной системы, когда режим дефицита сетевого интерфейса переходит в режим его профицита.

2. Проведен анализ выявленных режимов работы сетевого интерфейса в модульной многопроцессорной кластерной системе. Показано, что для получения высоких оценок эффективности и ускорения вычислений в кластерной системе необходимо, чтобы она функционировала в режиме дефицита сетевого интерфейса. Выявлено, что основная особенность режима

профицита сетевого интерфейса состоит в том, что коммутатор будет сталкиваться с перегрузками, когда сумма входящих трафиков превышает сумму выходящих. При этом основные характеристики эффективности такой кластерной системы будут существенно ухудшаться.

3. Определены условия работы коммутационной матрицы коммутатора в режиме сквозной коммутации (*cut-through*) с тем, чтобы информация передавалась без использования процедуры буферизации. Такой подход обеспечивает передачу пакетов с наибольшей скоростью, что приводит к улучшению оценок эффективности кластерной системы в целом.

4. Установлены оценки эффективности кластерной системы при организации многоканальных режимов функционирования сетевого интерфейса многопроцессорной кластерной системы.

5. Выведены аналитические соотношения для определения оптимального числа узлов многопроцессорной кластерной системы, когда соответствующая задача будет решаться за минимально возможное время.

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

Кроме того, в монографии проведен анализ исследования перспектив применения современных коммуникационных технологий в многопроцессорных кластерных системах. При этом основное внимание уделялось особенностям влияния сетевого интерфейса многопроцессорной вычислительной системы на оценки ее эффективности.

При этом:

1. Исследованы особенности формирования архитектуры сетевого интерфейса кластерной системы на основе применения технологий *Myrinet, Fibre Channel, 10Gb Ethernet*.

2. Для анализа согласования выбранного сетевого оборудования введены основные сетевые коэффициенты кластерной системы: коэффициент пропускной сети кластера, коэффициент пропускной способности

коммутатора, а также характеристика полосы пропускания коммутационной шины кластерной системы.

3. Выполнен сравнительный анализ оценок эффективности многопроцессорной кластерной системы для различного типа сетевых технологий. Показана перспективность применения технологии *Fibre Channel* при конструировании многопроцессорных кластерных систем в вузовских разработках.

4. Для увеличения пропускной способности сети кластера предложена и обоснована процедура "связывания каналов" (технология *channel bonding).* На основе анализа основных сетевых характеристик кластерной системы показана процедура выбора и согласования предложенного сетевого оборудования.

5. Выполнен сравнительный анализ результатов расчета основных оценок эффективности кластерной системы без реорганизации архитектуры сетевого интерфейса и после введения симметричных вычислительных сетей. Такие сети работают синхронно в результате реализации технологии *channel bonding.* Показано, что за счет повышения скорости обмена данными между узлами вычислительной системы удалось снизить загрузку каналов, которые соединяют эти узлы. Такой подход позволил существенно повысить оценки ускорения вычислений и значительно уменьшить время решения задачи.

6. Показано влияние ценового фактора современных коммуникационных технологий на эффективность распараллеливания и время решения задачи.

Качественный этап развития многопроцессорных кластерных систем лежит в области сопряжения нескольких модулей многопроцессорных систем в единый вычислительный комплекс для решения указанного типа задач.

При этом:

1. Раскрыты особенности сопряжения модульных многопроцессорных систем для расширения вычислительных возможностей при решении сильносвязанных задач.

2. Реализация реконфигурируемой сети позволила повысить эффективность кластерной системы, адаптируя структуру сети системы для решения каждого конкретного типа задач.

3. Введение подсетей загрузки системы, диагностики и обмена данными позволило разгрузить сети вычислительной системы повысить ее доступность и производительность.

4. Реализация промежуточного буфера памяти управляемых коммутаторов избавляет от необходимости синхронизации данных при сетевом обмене, когда реализуется процесс отправки и приема пакетов. При этом уменьшается загрузка *CPU*, что повышает эффективность и производительность кластерной системы в целом.

Литература

1. Пат. 57663 Україна, МПК G06F 15/16 (2011.01). Модуль високоефективної багатопроцесорної системи підвищеної готовності / Іващенко В.П., Башков Є.О., Швачич Г.Г., Ткач М.О.; власники: Національна металургійна академія України, Донецький національний технічний університет. – № u 2010 09341; заявл. 26.07.2010; опубл. 10.03.2011, Бюл. № 5.

2. Башков Є.О. Високопродуктивна багатопроцесорна система на базі персонального обчислювального кластера / Є.О. Башков, В.П. Іващенко, Г.Г. Швачич // Наукові праці Донецького національного технічного університету. Серія “Проблеми моделювання та автоматизації проектування”. – Вип. 9 (179). – Донецьк: ДонНТУ, 2011. – С.312 – 324.

3. Іващенко В.П. Дослідження оцінок ефективності модульної багатопроцесорної кластерної системи / В.П. Іващенко, Г.Г. Швачич, Є.О. Башков // Наукові праці Донецького національного технічного університету.

Серія “Інформатика, кібернетика та обчислювальна техніка”. – Вип. 13 (185). – Донецьк: ДонНТУ, 2011. – С. 33 – 43.

4. [Электронный ресурс]. – Режим доступа: http://tinklai.dkd.lt/administravimas/optinis_kabelis.htm#_Toc118999641.

5. [Электронный ресурс]. – Режим доступа: http://grid.imbg.org.ua/ Обчислювальний кластер ІМБГ НАН України.

6. [Электронный ресурс]. – Режим доступа: http://www.ixbt.com/cpu/clustering.shtml.

7. Fast Ethernet vs Myrinet выбор сетевой технологии для кластерной системы [Электронный ресурс]. – Режим доступа: http://www.nestor. minsk.by/sr/2000/07/00705.html.

8. Результаты тестирования вычислительного кластера [Электронный ресурс]. – Режим доступа: http://ias.csa.ru/CSA/MICRO/test-res/.

9. [Электронный ресурс]. – Режим доступа: http://www.top500.org.

10. 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.

11. Швачич Г.Г. Про проблему математичного моделювання термічної обробки довгомірного сталевого виробу / Г.Г. Швачич, М.О. Ткач // VII Intrenational Conference “Strategy of Quality in Industry and Education”, June, 3 – 10. – 2011, Varna; Bulgaria . – Proceedings. – V. 2. – P. 561 – 567.

12. Іващенко В.П. Дослідження оцінок ефективності модульної багатопроцесорної кластерної системи / В.П. Іващенко, Г.Г. Швачич, Є.О. Башков // Наукові праці Донецького національного технічного університету. Серія “Інформатика, кібернетика та обчислювальна техніка”. – Вип. 13 (185). – Донецьк: ДонНТУ, 2011. – С. 33 – 43.

13. Сбітнєв Ю.І. Дослідження оцінки ефективності багатопроцесорної кластерної системи / Ю.І. Сбітнєв, Г.Г. Швачич, М.О. Ткач // VI Intrenational

Conference “Strategy of Quality in Industry and Education”, June, 1 – 8 2010, Varna; Bulgaria . – Proceedings. – V. 2. – P. 288 – 296.

14. Информационные системы и технологии: монография. Кн. 3. / А.А. Белов, В.П. Иващенко, Е.А. Башков [и др.]. – Красноярск: Научно-инновационный центр, 2011. – 303 с.

15. [Электронный ресурс]. – Режим доступа: http://grouper.ieee.org/groups/ 802/3/ab/.

16. HP Networking [Электронный ресурс]. – Режим доступа: http://3com.com.

17. [Электронный ресурс]. – Режим доступа: http://ig.by/setevye-karty/setevaya-karta-3com-996b-t-server-adapter-pci-101001000.html

http://www.networkmanuals.ru/source/full.php?catalogue=3&sub=1149&cc=1025 &message=10.

18. Серия оборудования 3COM Super Starck3 [Электронный ресурс]. – Режим доступа: http://www.novacom.ru/products/3Com/SS3_Trans_guide.pdf.

19. [Электронный ресурс]. – Режим доступа: http://cluster.linux-ekb.info.

20. [Электронный ресурс]. – Режим доступа: http://kafvt.narod.ru/Osia/Glava4.htm.

21. Кластерные решения [Электронный ресурс]. – Режим доступа: http://www.hardline.ru/2/22/1559.

22. [Электронный ресурс]. – Режим доступа: http://www.myricom.com/.

23. Высокопроизводительный вычислительный кластер ВЦ РАН [Электронный ресурс]. – Режим доступа: http://www.ccas.ru/depart/kopytov/ot2003_3.htm.

24. Семенов Ю.А. Канальный протокол Fibre Channel [Электронный ресурс] / Ю.А. Семенов. – Режим доступа: http://book.itep.ru/4/41/f_ch 4112.htm.

25. Адаптер (HBA) Fibre Channel [Электронный ресурс]. – Режим доступа: http://www.rayton.ru/product/adapter-hba-fibre-channel-qlogic-sanblade-qla2462-ck/.

26. Масштабируемый коммутатор Fibre Channel -QLogic SANbox 5600Q [Электронный ресурс]. – Режим доступа: http://www.datasystems.ru/goods_SB5600Q-08A.htm#info.

27. [Электронный ресурс]. – Режим доступа: http://www.dscon.ru/san/qlogic_hba_QLE2562.htm.

28. [Электронный ресурс]. – Режим доступа: http://www.dell.com/us/business/p/brocade-300/pd.

29. 10 Gigabit Ethernet [Электронный ресурс]. – Режим доступа: http://kunegin.narod.ru/ref1/giga/10giga.htm.

30. [Электронный ресурс]. – Режим доступа: http://www.ecolan.ru/imp_info/standarts/change/tengig/.

31. [Электронный ресурс]. – Режим доступа: http://activka.ua/lntel-ethernet-server-adapter-x520-t2.html.

32. [Электронный ресурс]. – Режим доступа: http://www.hpcwire.com/hpcwire/2008-11-18/arista_announces_10gbe_ecosystem.html.

Научное издание

Иващенко Валерий Петрович,
Башков Евгений Александрович,
Швачич Геннадий Григорьевич,
Ткач Максим Александрович

СОВРЕМЕННЫЕ КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ В МОДУЛЬНЫХ МНОГОПРОЦЕССОРНЫХ СИСТЕМАХ: ОПЫТ ИСПОЛЬЗОВАНИЯ, ИССЛЕДОВАНИЕ ОЦЕНОК ЭФФЕКТИВНОСТИ, ПЕРСПЕКТИВЫ ПРИМЕНЕНИЯ

Монография

Редактор
Корректор
Компьютерная верстка

Подписано в печать
Формат
Бумага типографская. Заказ №
Тираж 500 экз.

Издательство