

УДК 004.272.43.003.13

Е.А. Башков<sup>1</sup>, В.П. Иващенко<sup>2</sup>, Г.Г. Швачич<sup>2</sup>, М.А. Ткач<sup>2</sup>  
<sup>1</sup>Донецкий национальный технический университет,  
<sup>2</sup>Национальная металлургическая академия Украины  
[bashkov@pmi.dgtu.donetsk.ua](mailto:bashkov@pmi.dgtu.donetsk.ua)

## Перспективы применения современных коммуникационных технологий и исследование их влияния на эффективность многопроцессорных кластерных систем

Статья посвящена 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.

**Персональный вычислительный кластер, управляемые коммутаторы, реконфигурируемая сеть, вычислительные узлы, ускорение, эффективность**

### Введение

Практически одновременно с появлением первых многопроцессорных кластерных 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 реализации сетевого интерфейса многопроцессорной системы. Неудачная реализация сетевого интерфейса порой может свести на нет эффект от увеличения числа используемых процессоров. 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 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 распространенных можно отнести следующие: *Fast Ethernet*, *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*. 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*.

Так, на основе сетевой технологии *Ethernet* было реализовано “блейд” серверное решение многопроцессорной системы, при которой несколько однотипных материнских модулей устанавливаются в одном корпусе [1]. Основные особенности конструирования ее архитектуры изложены в [2]. В конфигурации кластера было избрано шесть лезвий и модульный принцип его реализации. Это обеспечивает, в случае необходимости, его расширение за счет установки дополнительных модулей. Каждый узел работает под управлением собственной копии операционной системы. Состав и мощность узлов такого кластера могут меняться, что позволяет создавать неоднородные системы. Коммутирующая сеть соединяет процессоры друг с другом. 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. В работе [3] достаточно детально изложены вопросы исследования эффективности такой многопроцессорной вычислительной системы.

Ряд исследований посвящен тестированию различных сетевых технологий [4, 5], 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 перспектив применения современных коммуникационных технологий в многопроцессорных кластерных системах. Основное внимание уделяется влиянию сетевого интерфейса на оценки эффективности кластерной системы.

### **Постановка проблемы исследований, ее актуальность**

В настоящее время появилась уникальная возможность создавать недорогие установки суперкомпьютерного типа – многопроцессорные вычислительные кластеры.

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.

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

Ясно и другое, что с помощью 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. При этом эффективность распараллеливания вычислений зависит от многих факторов, однако одним из определяющих является выбор и организация сетевого интерфейса. 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 процессоров. В этой связи, чем выше будет пропускная способность вычислительной сети кластера, тем быстрее будут считаться пользовательские параллельные задачи,

запущенные 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.

### **Анализ последних исследований и публикаций**

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

Достаточно полный и всесторонний анализ оценок эффективности модульной многопроцессорной кластерной системы приведен в [3, 6].

На первом этапе исследований рассмотрим особенности формирования архитектуры сетевого интерфейса кластерной системы [1, 2] и основные режимы его работы. Для удобства освещения такого подхода предварительно отметим, что вычислительная сеть кластерной системы имеет две основные характеристики – пропускная 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 накладных расходов, возникающих 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 данными 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, 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.

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

**Сетевые кабели.** Для сетевых соединений в многопроцессорной вычислительной системе применялась технология *GigabitEthernet* [7]. При этом технология *100BASE-T*, *IEEE 802.3ab* – стандарт, использующий витую пару категорий *5e*. 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. Например, можно привести следующие: *Compx*, *Intel*, *Macronix* и др. Однако для данной кластерной системы предполагалось использование технологии связывания каналов (*channel bonding*), которая позволяет объединять несколько сетевых адаптеров в один скоростной виртуальный канал. В этой связи, предпочтение было отдано производителю *3COM* [8]. Сетевые карты фирмы *3 COM*, 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.** 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.** Снижает 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 *3COM 3C996B-T* [9].

**Коммутаторы.** Одним из важных элементов сетевого интерфейса кластерной системы являются устройства коммутации сетевых каналов.

При выборе коммутирующих устройств, в первую очередь, необходимо учитывать возможность использования технологии связывания каналов *channel bonding*. Проблема связывания каналов заключается в том, что при наличии технологии *channel bonding* появляются две или несколько сетевых карт с одинаковым *MAC*-адресом. В обычном режиме работы коммутатор будет интенсивно перестраивать свои внутренние таблицы портов, переназначая *MAC*-адрес с одного порта на другой. Это может привести либо к полной неработоспособности канала, либо к значительным потерям пакетов и существенному снижению производительности сети. Для обеспечения нормальной работы таких связанных каналов в коммутаторе должны быть предусмотрены функции *Link Aggregation* или, поддержка коммутатором стандарта *IEEE 802.3ad*.

Принимая во внимание отмеченное, был выбран восьмипортовый коммутатор *SuperStak II Switch 9000* производителя *3COM LG* [10], который имеет пропускную способность 17,2 Гбит/с. Это означает, что каждый из восьми портов может одновременно с другими в дуплексном режиме передавать и принимать данные с суммарной скоростью 2 Гбит/с.

Технические характеристики сетевого оборудования освещаемой многопроцессорной кластерной системы приведены в табл. 1.

Таблица 1. Технические характеристики сетевого оборудования кластерной системы

Сетевой кабель	Тип	<i>GigabitEthernet</i>
	Пропускная способность	1 Гбит/с
	Стандарт	<i>5e</i>
	Цена	\$ 6
Сетевой адаптер	Тип	<i>3C996B-T</i>
	Производитель	<i>3COM</i>
	Пропускная способность	1 Гбит/с
	Цена	\$ 60
Коммутатор	Тип	<i>SuperStak II Switch 9000</i>
	Производитель	<i>3COM</i>
	Пропускная способность	17,2 Гб/с
	Цена	\$ 900

Итак, можно отметить, что ориентировочная цена сетевого оборудования будет составлять около 1300 у.е.

При указанных сетевых характеристиках кластерной системы были проведены соответствующие вычислительные эксперименты, а также исследованы характеристики эффективности многопроцессорной системы. Исходные данные для этого режима исследований представлены в табл. 2.

Таблица 2. Исходные данные для расчета характеристик эффективности многопроцессорной системы

$V_p$	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. 3.

Таблица 3. 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	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

Проведенный анализ полученных результатов моделирования показал следующее. Наибольшая величина ускорения вычислений достигается на шести узлах многопроцессорной системы и соответствует значению, равному 3,25. Время счета уменьшается со 100 с до 30,81 с.

Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы представлена на рис. 1.

В работе [2] для увеличения пропускной способности сети кластера применялась процедура "связывания каналов" или технология *channel bonding*. Технология связывания каналов (*channel bonding*) позволяет объединять несколько сетевых адаптеров в один скоростной канал. Схема организации сетевого интерфейса по технологии *channel bonding* приведена на рис. 2.

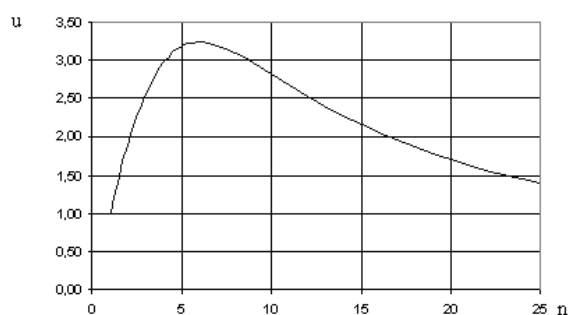
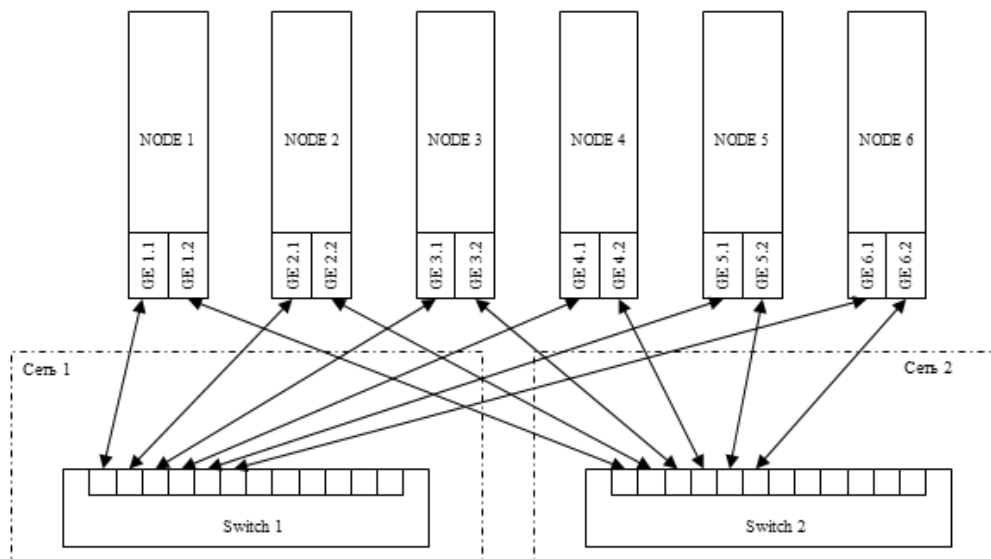


Рисунок 1 – Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы

Особенность организации такого сетевого интерфейса (рис. 2) состоит в том, что здесь сконфигурированы две симметрично работающие вычислительные сети (сеть 1 и сеть 2) на основе двух коммутационных матриц в сети обмена данными (коммутаторы *Switch 1* и *Switch 2*). Приведенная схема организации сетевого интерфейса включает на каждом вычислительном узле две однотипные сетевые карты (3C996B-T) и два однотипных коммутатора (*SuperStak II Switch 9000*). Для конфигурации приведенных сетевых интерфейсов выполняются основные операции по настройке режима *Link Aggregation*.

Благодаря такому подходу (рис. 2), появляется возможность, с одной стороны, организовать равномерное распределение нагрузки (приема/передачи данных) между соответствующими узлами кластерной системы, а с другой – 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.

Рисунок 2 – Схема організації мережевого інтерфейса по технології *channel bonding*

Вообще заметим, что технология *channel bonding* может порождать некоторые проблемы, связанные с выбором коммутаторов и их настройки. Особенности выбора сетевого интерфейса для использования технологии *channel bonding* достаточно детально освещаются в [11].

Исходные данные для расчета основных характеристик многопроцессорной кластерной системы с использованием технологии *channel bonding* перечислены в табл. 2.

Была проведена процедура расчета основных характеристик эффективности многопроцессорной кластерной системы с использованием технологии *channel bonding*. Полученные результаты моделирования сведены в табл. 4.

Для наглядности приведем график ускорения вычислений при указанных условиях эксперимента.

Сравнительный анализ результатов расчета без реорганизации архитектуры сетевого интерфейса (табл. 3) и после введения симметричных вычислительных подсетей, которые работают одновременно в результате реализации технологии *channel bonding* (табл. 4), показал, 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, а время решения задачи снизилось с 30,81 с до 22,4 с.

Таблица 4. Результаты расчета основных характеристик многопроцессорной кластерной системы с использованием технологии *channel bonding*

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
16	6,25	21,21	27,46	3,64	0,23

Итак, проведенный обзор показывает, что, с одной стороны, вопросам исследования сетевого интерфейса многопроцессорных кластерных систем уделяется достаточно серьезное внимание. С другой стороны, такие исследования охватывают самые разнообразные режимы функционирования кластерных многопроцессорных систем.

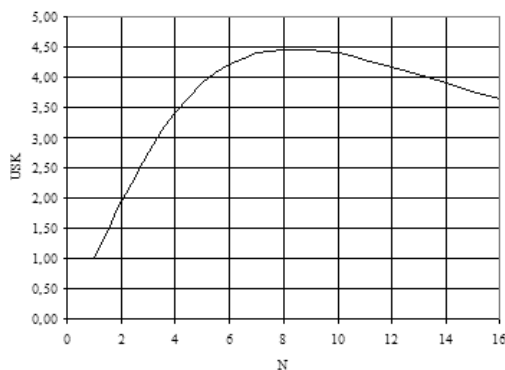


Рисунок 3 – Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы при использовании технологии *channel bonding*

Однако можно отметить, что на сегодняшний день проблема влияния современных коммуникационных технологий кластерной системы на ее эффективность раскрыта не полностью. Проведенные в данной работе исследования направлены на устранение указанного недостатка и развивают подход к исследованию модульных многопроцессорных систем, освещенный авторами в [2, 3, 6].

### Цель и задачи исследований

Основная цель исследований, представленных в данной работе, заключается в анализе исследования перспектив применения современных коммуникационных технологий в многопроцессорных кластерных системах.

При этом основное внимание уделяется особенностям влияния сетевого интерфейса такой системы на оценки ее эффективности.

В результате проведенных исследований 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 процедуру "связывания каналов" (технология *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.

### Изложение основного материала исследований

#### Сетевая технология *Myrinet*

На первом этапе исследований рассмотрим особенности формирования архитектуры сетевого интерфейса кластерной системы на основе применения технологии *Myrinet*, предлагаемой компанией *Myricom* [12].

В списке *Top500* 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 [13].

#### Сетевая технология *Fibre Channel*

*Fibre Channel (FC)* (англ. *fibre channel* – волоконный канал) – тип коммуникационной среды для высокоскоростной передачи данных [14]. Благодаря высокой скорости передачи данных, малой задержке и расширяемости практически не имеет аналогов в своей области. В последние годы область применения сетевой технологии *FC* постепенно перемещается в сегмент конструирования высокопроизводительных вычислительных систем.

Принимая во внимание отмеченное, рассмотрим особенности выбора сетевого интерфейса на основании 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. На первом этапе его развития казалось, что этот интерфейс будет доминирующим. Однако, по мере повышения быстродействия *Ethernet*, он перешел на второй план. А за последние годы интерфейс *FC* обрел второе дыхание. Произошло это благодаря симбиозу *Ethernet u Fibre Channel - FCoE (Fibre Channel over Ethernet)*. Для конструирования многопроцессорных вычислительных систем рекомендуется использовать технологию *FC* с пропускной способностью  $V_s = 4$  Гбит/с. Такой сетевой интерфейс удачно себя зарекомендовал именно

для конструювання модульних багатопроцесорних систем.

*Сетевий адаптер.* В качестве сетевых адаптеров необходимо использовать карты, поддерживающие работу в стандарте *FC*. Однако при их выборе необходимо принять во внимание то обстоятельство, что для конструируемой кластерной системы предполагается использование технологии связывания каналов (*channel bonding*), которая позволяет объединять несколько сетевых адаптеров в один скоростной виртуальный канал. В этой связи, предпочтение было отдано адаптеру *QLA2460* [15] производителя *QLogic*. Такой адаптер обеспечивает высокие показатели производительности и доступности данных, а также предоставляет набор интеллектуальных сетевых средств, необходимых при создании модульных многопроцесорных вычислительных систем.

Сетевой адаптер *QLA2460* является высокопроизводительным и наиболее успешным в своем классе. Существенным отличием адаптера *QLA2460* 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 многопроцесорных 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 коммутатор *QLogic SANbox 5600Q* [16]. Такой коммутатор обеспечивают естественную масштабируемость и высокую производительность на основе объединяющего шасси в расширяемом, легко администрируемом решении. Кроме того, коммутатор *QLogic SANbox 5600Q* 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.

Технические характеристики сетевого оборудования предполагаемой многопроцесорной кластерной системы приведены в табл. 5.

Таблица 5. Технические характеристики сетевого оборудования кластерной системы

Сетевой кабель	Тип	<i>Fibre Channel</i>
	Пропускная способность	4 Гбит/с
	Стандарт	<i>FC</i>
	Цена	\$ 50
Сетевой адаптер	Тип	<i>QLA2460</i>
	Производитель	<i>QLogic</i>
	Пропускная способность	4 Гбит/с
	Цена	\$ 1100
Коммутатор	Тип	<i>QLogic SANbox 5600Q</i>
	Производитель	<i>QLogic</i>
	Пропускная способность	136 Гб/с
	Цена	\$ 2000

Итак, можно отметить, что ориентировочная цена сетевого оборудования будет составлять около 16000 у.е.

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

На первом этапе исследований рассмотрим проблему согласования устройств сетевого интерфейса.

Для этого необходимо сравнивать общую пропускную способность сети кластера

$$V_s = V_p \cdot N \quad (1)$$

и пропускную способность коммутатора ( $V_b$ ).

Для дальнейшего анализа сетевого интерфейса кластерной системы введем в рассмотрение коэффициент  $k_s$  и будем его трактовать, как коэффициент пропускной сети кластера, который определим следующим образом:

$$k_s = \frac{V_s}{V_b} \quad (2)$$

С учетом формулы (1), получаем:

$$k_s = \frac{V_p \cdot N}{V_b}. \quad (3)$$

Для развития предлагаемого подхода введем в рассмотрение понятие коэффициента пропускной способности коммутатора ( $k_b$ ):

$$k_b = \frac{V_b}{V_p \cdot N}. \quad (4)$$

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. 6.

Таблица 6. Исходные данные для расчета сетевых характеристик кластерной системы

$V_p$	4 Гбит/с
$V_b$	136 Гбит/с
$k$	1

Здесь  $k$  – количество симметричных вычислительных сетей в архитектуре кластерной системы.

Полученные результаты моделирования сведены в табл. 7.

Таблица 7. Результаты расчета основных сетевых коэффициентов кластерной системы

колич, узлов $N$	$k_s$	$K_b$	$c_k$
1	0,03	34,00	136,00
2	0,06	17,00	68,00
3	0,09	11,33	45,33
4	0,12	8,50	34,00
5	0,15	6,80	27,20
6	0,18	5,67	22,67
7	0,21	4,86	19,43
8	0,24	4,25	17,00
9	0,26	3,78	15,11
10	0,29	3,40	13,60
11	0,32	3,09	12,36
12	0,35	2,83	11,33
13	0,38	2,62	10,46
14	0,41	2,43	9,71
15	0,44	2,27	9,07
16	0,47	2,13	8,50
17	0,50	2,00	8,00
18	0,53	1,89	7,56
19	0,56	1,79	7,16

20	0,59	1,70	6,80
21	0,62	1,62	6,48
22	0,65	1,55	6,18
23	0,68	1,48	5,91
24	0,71	1,42	5,67
25	0,74	1,36	5,44
26	0,76	1,31	5,23
27	0,79	1,26	5,04
28	0,82	1,21	4,86
29	0,85	1,17	4,69
30	0,88	1,13	4,53
31	0,91	1,10	4,39
32	0,94	1,06	4,25
33	0,97	1,03	4,12
34	1,00	1,00	4,00
35	1,03	0,97	3,89

Проведем предварительный анализ полученных результатов. Очевидно, что с увеличением числа узлов кластерной системы будет возрастать пропускная способность сети кластера ( $V_s$ , формула 1). Тогда изменение коэффициента пропускной способности сети кластера ( $k_s$ , формула 3) будет осуществляться по линейному закону. С другой стороны, увеличение объема данных, пересылаемых между узлами кластера, приведет к тому, что нагрузка коммутатора будет возрастать, и его коэффициент пропускной способности ( $k_b$ ) будет уменьшаться по нелинейному закону (формула 4). 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 вида:

$$c_k = \frac{V_b}{N}. \quad (5)$$

В табл. 8 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 несколько превышать протокольные возможности сети. Однако можно отметить, что коммутаторы с меньшей пропускной способностью находятся в той же ценовой категории, что и предложенный. По этой причине дальнейшие исследования будут проведены с учетом выбранного сетевого оборудования.

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

Расчетные формулы для исследования таких характеристик согласуются с аналитическими соотношениями, принятыми в работе [3].

Исходные данные для этого режима исследований представлены в табл. 8, полученные результаты моделирования сведены в табл. 9.

Таблица 8. Исходные данные для расчета характеристик эффективности многопроцессорной системы

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

Таблица 9. 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

Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы показана на рис. 4.

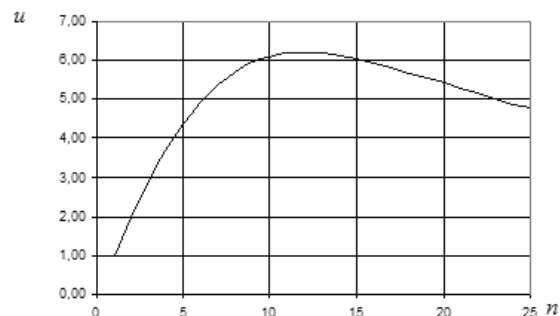


Рисунок 4 – Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы

Проведенный анализ полученных результатов моделирования показал следующее. Наибольшая величина ускорения вычислений при предложенном сетевом интерфейсе соответствует значению, равному 6,21. Время счета задачи уменьшается со 100 с до 16,11 с.

Далее заметим, что элементы сетевого интерфейса выбирались с учетом поддержки технологии *channel bonding*. Такой подход позволяет улучшить оценки эффективности модульной кластерной системы. Основные сетевые характеристики рассматриваемого режима работы кластерной системы приведены в табл. 10.

Таблица 10. Основные сетевые характеристики кластерной системы с использованием технологии *channel bonding*

$V_p$	4 Гбит/с
$V_b$	136 Гбит/с
$k$	2
$k_m$	2

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

Можно отметить, что ориентировочная цена предлагаемого комплекта сетевого оборудования будет составлять около 30000 у.е. Здесь стоимость сетевого оборудования определялась с учетом выбора двухпортового сетевого адаптера *QLE2562-CK*.

Моделирование основных сетевых характеристик в соответствии с формулами (3) – (5) показало, что, при равных прочих условиях, за счет изменения архитектуры сетевого интерфейса многопроцессорной системы будет расширяться

только полоса 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 выполнения необходимых расчетов, существенно улучшая характеристики эффективности и быстродействия.

Был проведен этап моделирования таких характеристик. Полученные результаты моделирования сведены в табл. 11.

Проведенный анализ полученных результатов моделирования показал следующее. Наибольшая величина ускорения вычислений при предложенном сетевом интерфейсе соответствует значению, равному 8,67. Время счета задачи уменьшается до 11,54 с.

Таблица 11. Результаты расчета основных характеристик эффективности многопроцессорной системы с использованием технологии *channel bonding*

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

Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы показана на рис. 5.

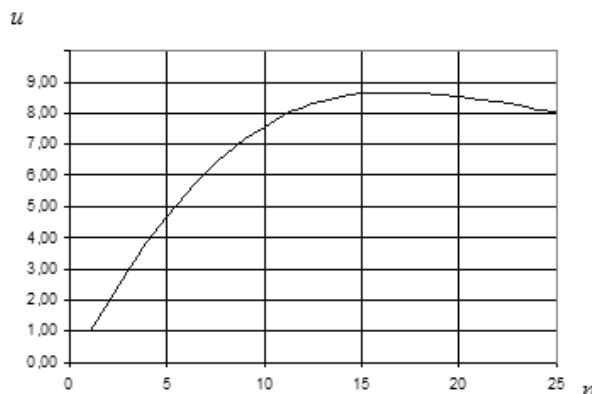


Рисунок 5 – Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы с использованием технологии *channel bonding*

#### Сетевая технология 10Gb Ethernet

Сетевая технология 10 *Gigabit Ethernet* или *10GbE* является новейшим и самым быстрым из существующих стандартов *Ethernet* [17]. 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 802.3an (10GBaseT)* 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, "посаженым" 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 работе и будет исследован сетевой интерфейс на базе технологии *10Gb Ethernet*.

Вообще заметим, что в настоящее время активно ведется разработка стандартов и 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.

Был проведен детальный анализ существующего сетевого оборудования для рассматриваемого стандарта, технические характеристики такого оборудования в предполагаемой многопроцессорной кластерной системы приведены в табл. 12.

Таблица 12. Технические характеристики сетевого оборудования кластерной системы

Сетевой кабель	Тип	<i>10Gb Ethernet</i>
	Пропускная способность	10 Гбит/с
	Стандарт	<i>10Gb Ethernet</i>
	Цена	\$ 10
Сетевой адаптер	Тип	<i>CNA Brocade 1010</i>
	Производитель	<i>Brocade</i>
	Пропускная способность	10 Гбит/с
	Цена	\$ 1100
Коммутатор	Тип	<i>ES 5508</i>
	Производитель	<i>Edge-Core</i>
	Пропускная способность	160 Гб/с
	Цена	\$ 8000

Таблица 13. Оценки влияния ценового фактора современных коммуникационных технологий на эффективность распараллеливания и время решения задачи

технология	колич. узлов, <i>N</i>	скорость сети, Гб/с	время счета, с	ускорение	цена сетевого интерфейса, у.е.
<i>GE</i>	6	1	30,81	3,25	1300
<i>GE+CB</i>	8	2	22,4	4,46	2900
<i>FC</i>	12	4	16,11	6,21	16000
<i>FC+CB</i>	16	8	11,54	8,67	30000
<i>10GbE</i>	18	10	10,35	9,66	31000
<i>10GbE+CB</i>	24	20	7,38	13,55	70000

Итак, можно отметить, что ориентировочная цена сетевого оборудования будет составлять около 31000 у.е.

В соответствии с ранее освещенной методикой, при указанных сетевых характеристиках кластерной системы были проведены соответствующие вычислительные эксперименты для согласования узлов сетевого интерфейса, а также исследованы характеристики эффективности многопроцессорной системы.

Проведенный анализ полученных результатов моделирования показал следующее. Наибольшая величина ускорения вычислений при предложенном сетевом интерфейсе соответствует значению, равному 9,66. Время счета задачи уменьшается со 100 с до 10,35 с.

Затем были проведены вычислительные эксперименты с учетом реорганизации архитектуры сетевого интерфейса для введения симметричных вычислительных сетей, которые работают одновременно в результате реализации технологии *channel bonding*. Ориентировочная цена такого комплекта сетевого оборудования будет составлять около 70000 у.е. Здесь стоимость сетевого оборудования определялась с учетом выбора двухпортового сетевого адаптера *CNA Brocade 1020*.

Такой подход позволил существенно повысить оценки ускорения вычислений (13,55), а время решения задачи снизилось с 10,35 с до 7,38 с.

### Анализ полученных результатов

Анализ исследования 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. 13 приведены оценки влияния ценового фактора современных коммуникационных технологий на эффективность распараллеливания и время решения задачи.

Анализ табл. 13 показывает перспективность применения технологии *Fibre Channel* при конструировании многопроцессорных кластерных систем в вузовских разработках.

### **Выводы и перспективы последующих исследований**

В работе 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

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 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 существенно уменьшить время решения задачи.

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

### **Список литературы**

1. Модуль високоефективної багато процесорної системи підвищеної готовності: пат. 57663 Україна, МПК G06F 15/16 (2011.01) / 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 багато процесорна 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. – 2011. – Вип. 9(179). – С.312 - 324.

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

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

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

6. Сбитнев Ю.І. Дослідження оцінки ефективності багато процесорної кластерної системи / Ю.І. Сбитнев, Г.Г. Швачич, М.О. Ткач // VI International Conference ["Strategy of Quality in Industry and Education"], 1 - 8 June 2010, Varna; Bulgaria. – Proceedings. – V. 2. – P. 288 – 296.

7. [Электронный ресурс]. – Режим доступа: <http://groupier.ieee.org/groups/802/3/ab/>.

8. HP Networking [Электронный ресурс]. – Режим доступа: <http://3com.com>.

9. NetworkManuals.Ru [Электронный ресурс]. – Режим доступа: <http://www.networkmanuals.ru/source/full.php?catalogue=3&sub=1149&cc=1025&message=10>.

10. Серія обладнання 3COM Super Starck3 [Электронный ресурс]. – Режим доступа: [http://www.novacom.ru/products/3Com/SS3\\_Trans\\_guide.pdf](http://www.novacom.ru/products/3Com/SS3_Trans_guide.pdf).

11. Практическое руководство по параллельным вычислениям [Электронный ресурс]. – Режим доступа: <http://cluster.Linux-ekb.info>.
12. [Электронный ресурс]. – Режим доступа: <http://www.myricom.com/>.
13. Высокопроизводительный вычислительный кластер ВЦ РАН [Электронный ресурс]. – Режим доступа: [http://www.ccas.ru/depart/kopytov/ot2003\\_3.htm](http://www.ccas.ru/depart/kopytov/ot2003_3.htm).
14. Семенов Ю.А. Канальный протокол Fibre Channel [Электронный ресурс] / Ю.А. Семенов. – Режим доступа: [http://book.iter.ru/4/41/f\\_ch\\_4112.htm](http://book.iter.ru/4/41/f_ch_4112.htm).
15. Адаптер (HBA) Fibre Channel [Электронный ресурс]. – Режим доступа: <http://www.rayton.ru/product/adapter-hba-fibre-channel-qlogic-sanblade-qla2462-ck/>
16. Масштабируемый коммутатор Fibre Channel -QLogic SANbox 5600Q [Электронный ресурс]. – Режим доступа: [http://www.datasystems.ru/goods\\_SB5600Q-08A.htm#info](http://www.datasystems.ru/goods_SB5600Q-08A.htm#info).
17. 10 Gigabit Ethernet [Электронный ресурс]. – Режим доступа: <http://kunegin.narod.ru/ref1/giga/10giga.htm>.

Надійшла до редколегії 15.10.2011

**Є.О. БАШКОВ, В.П. ІВАЩЕНКО,  
Г.Г. ШВАЧИЧ, М.О. ТКАЧ**

Донецький національний технічний університет  
Національна металургійна академія України

**E.A. BASHKOV, V.P. IVASHENKO,  
G.G. SHVACHYCH, M. A. TKACH**

Donetsk national technical university  
National metallurgical academy of Ukraine

**Перспективи застосування сучасних  
комунікаційних технологій та дослідження їх  
впливу на ефективність багатопроцесорних  
кластерних систем**

**Prospects for Application of Modern  
Communication Technologies and Research of their  
Influence on Efficiency of Multiprocessor Cluster  
Systems**

Статтю присвячено проблемі дослідження перспектив застосування сучасних комунікаційних технологій в багатопроцесорних кластерних системах. Основна увага приділяється впливу мережевого інтерфейсу на оцінки ефективності кластерної системи. Розкриті питання узгодження мережевого інтерфейсу з обчислювальними вузлами в багатопроцесорній системі.

The article is devoted to the problem of studying the prospects for application of modern communication technologies in multiprocessor cluster systems. The main attention is focused on the influence of network interface on estimations of the cluster system efficiency. The questions of matching the network interface with computational nodes in the multiprocessor system are exposed in the present paper.

*персональний обчислювальний кластер,  
керовані комутатори, реконфігурована мережа,  
обчислювальні вузли, прискорення,  
ефективність*

*personal calculable cluster, guided switchboards,  
reconfigurable network, calculable knots,  
acceleration, efficiency*