



УДК 004.272.45

# МЕТОД ОДНОРОДНОГО РАСШИРЕНИЯ СЕТЕЙ СВЯЗИ МНОГОПРОЦЕССОРНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

В. С. Подлазов, В. В. Соколов

*Институт проблем управления им В. А. Трапезникова РАН, г. Москва*

Предложен метод однородного расширения любых системных сетей, который сохраняет такие их функциональные свойства, как способ и качество маршрутизации и задержки передачи данных, и позволяет собирать параллельным образом большие системные сети из малых сетей с заданными функциональными свойствами.

*Светлой памяти  
Ивери Варламовича Прангишвили  
посвящается*

## ВВЕДЕНИЕ

В середине 1970-х г. И. В. Прангишвили с сотрудниками предложили концепцию однородных параллельных вычислительных систем [1], которая привела к созданию, в частности, многопроцессорной вычислительной системы (МВС) ПС2000. Межпроцессорной сетью связи в этой МВС был кольцевой коммутатор, который предоставлял некоторые возможности для параллельной передачи данных. В работах [2–11] эти возможности были в максимальной степени расширены благодаря применению мультикольцевых сетей связи, которые оказались равносильными сетям связи со структурой обобщенных гиперкубов. Однако сам метод распараллеливания сети связи оставался узкоспециализированным — ориентированным на мультикольцевые или гиперкубовые структуры. Вопрос о существовании единого метода распараллеливания сети связи с произвольной структурой в литературе даже не ставился. В данной статье этот вопрос рассматривается в общем виде и делается первый шаг к его решению. Этот шаг основывается на описании сетей связи с использованием структур, называемых неполными уравновешенными блок-схемами (УБС) [12]. Математическая теория УБС была развита в комбинаторике в середине прошлого века для целей планирования сложных экспериментов в селекции, генетике и др. В настоящей статье она получает новое применение для описания сетей связи МВС. В настоящее время за этими сетями утвердился термин системные сети (System Area Network — SAN) [13], который и используется в дальнейшем.

## 1. МУЛЬТИКОЛЬЦЕВЫЕ СИСТЕМНЫЕ СЕТИ

В качестве отправной точки рассмотрим кольцевой коммутатор МВС ПС2000, в котором параллельно могут осуществляться некоторые передачи одинаковых по длине кадров. Теоретическая модель кольцевого коммутатора выглядит следующим образом [14]. Кольцо используется в полудуплексном режиме, при котором все параллельные передачи осуществляются в одну сторону с возможностью смены направления передачи после каждого такта. Каждый такт равен длительности кадра. За один такт может осуществляться параллельная передача кадров по неперекрывающимся маршрутам.

В этой модели произвольная перестановка кадров между узлами осуществляется за  $N - 1$  тактов, где  $N$  — число узлов (процессоров) на кольце [14]. При перестановке имеет место однозначная адресация кадров «из одного узла в один и только один узел».

В более общем случае произвольной индивидуальной адресации кадров можно говорить только о среднем числе кадров, переданных за один такт — емкости сети  $C$ . В симплексном (с однонаправленной передачей) кольце при равновероятной адресации кадров она равна 2, в полудуплексном кольце (коммутатор ПС2000) при передаче кадров по кратчайшим маршрутам — 4, а в случае дуплексной кольцевой сети — 8 [2, 3, 5, 6]. Для достижения последнего значения необходимо составить дуплексное кольцо из двух симплексных колец со встречными направлениями передачи.

Четырехкратное увеличение пропускной способности при переходе от симплекса к дуплексу является коммулятивным эффектом от увеличения степени параллелизма сети по двум направлениям — за счет увеличения



числа колец и за счет увеличения числа параллельно передаваемых кадров в каждом кольце. Хотелось бы сохранить этот нелинейный рост пропускной способности и при дальнейшем увеличении числа колец.

Это удастся сделать, если использовать кольца с разными шагами. Обычное однонаправленное кольцо имеет шаг 1. Кольцо с обратным направлением передачи (встречное кольцо) имеет шаг  $N - 1$  или  $-1$  по  $\text{mod } N$ . Набор колец с разными шагами образует мультикольцо. На рис. 1 представлено мультикольцо с девятью узлами, состоящее из колец с шагами 1, 2, 3 и 6 (или  $-3$ ). Для мультикольца решалась задача поиска такой его структуры, которая обеспечивает бесконфликтную реализацию произвольной перестановки за минимальное число тактов, а при произвольной адресации обеспечивает максимальную емкость мультикольца.

Для обоих видов адресации рассматривалось коммутируемое мультикольцо [3, 4, 7, 8], узел которого содержит однокаскадный коммутатор, осуществляющий перестановку кадров между кольцами. Для произвольной адресации дополнительно рассматривалось некоммутируемое мультикольцо, в котором любой кадр доставляется от источника к приемнику по тому кольцу, в которое он был передан [3, 5 и 6]. Кадры передаются по кольцу с кратчайшим маршрутом. В некоммутируемом мультикольце достигаются минимальные времена передачи кадров.

Для обоих видов мультиколец поставленные задачи были решены, но с разной степенью общности. Для некоммутируемых мультиколец при произвольной адресации все результаты носят эмпирический характер, т. е. они получены в результате стохастической оптимизации структуры мультикольца для каждого заданного числа узлов [3, 5, 6]. Выяснилось, что емкость мультиколец с оптимальной структурой (набором колец) растет прямо пропорционально квадрату числа колец в них. Это оз-

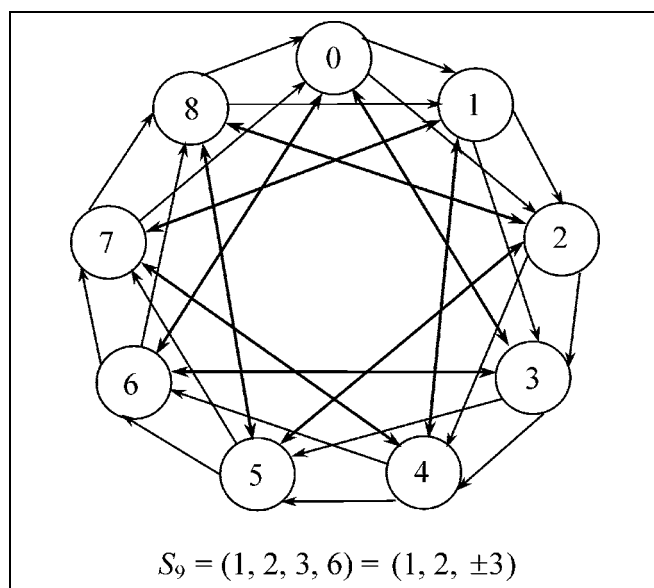


Рис. 1. Троичное мультикольцо с девятью узлами

Таблица 1

Наращиваемые мультикольца, для которых  $C \approx N$

| Число узлов | Число колец | Набор колец  |
|-------------|-------------|--|
| 16          | 4           | $\pm 1, \pm 3$   |
| 32          | 6           | $\pm 1, \pm 2, \pm 3$                                      |
| 64          | 8           | $\pm 1, \pm 2, \pm 3, \pm 7$                               |
| 128         | 12          | $\pm 1, \pm 2, \pm 3, \pm 5, \pm 7, \pm 9$                 |
| 256         | 16          | $\pm 1, \pm 2, \pm 3, \pm 5, \pm 7, \pm 9, \pm 11, \pm 13$ |

начает, что при числе колец  $\sim \sqrt{N}$  емкость мультикольца становится прямо пропорциональной числу узлов  $N$ . Были построены наращиваемые мультикольца с такими свойствами для  $N \leq 256$ . Частично они представлены в табл. 1. При этом остается неясным, сохраняются ли эти свойства для больших  $N$ , так как строгого обоснования для них найдено не было.

Для коммутируемых мультиколец была построена теория реализации произвольной перестановки при пакетной коммутации по статическим расписаниям. Эта теория построена для  $p$ -ичных мультиколец, шаги колец которых задаются цифрами  $p$ -ичной системы счисления [4, 7–11]. На рис. 1 представлено троичное мультикольцо. Эта теория справедлива и для обобщенных  $p$ -ичных гиперкубов. В ней были выяснены условия бесконфликтной реализации произвольных перестановок при пакетной коммутации и достигаемые при этом сетевые характеристики.

В частности,  $p$ -ичное мультикольцо с  $N = p^r$  узлами, содержащее  $m = (p - 1)r$  колец, позволяет бесконфликтно осуществить произвольную перестановку за  $n = 2 \left\lceil \frac{2(\sqrt{N} - 1)}{(p - 1)r} \right\rceil$  тактов при четном  $r$ . При  $p = \sqrt{N}$  мультикольцо превращается в распределенный однокаскадный неблокируемый самомаршрутизируемый коммутатор.

В случае произвольной адресации  $p$ -ичное мультикольцо с  $N = p^r$  узлами имеет емкость  $C = p(N - 1) = N^{1/r}(N - 1)$  [3, 4]. При  $p = 2$  и  $r = \log^2 N$  имеем  $C = 2(N - 1)$ , а при  $p = \sqrt{N}$  и  $r = 2 - C \approx N^{3/2}$ .

Здесь обращает на себя внимание то, что при одной схеме распараллеливания системной сети (в виде мультикольца) для разных видов мультиколец пришлось использовать разные схемы описания и получить разные по степени общности результаты. Хотелось бы найти такой метод распараллеливания системных сетей, который имел бы одинаковую схему описания для любой сети. По-видимому, такой метод удалось найти. Он определяется теорией уравновешенных блок-схем с некоторыми нашими дополнениями.

## 2. УРАВНОВЕШЕННЫЕ БЛОК-СХЕМЫ

Что представляют собой блок-схемы? Для их определения используются только три понятия: элемент, блок и вхождение элемента в блок [12]. Параметрами являются: число элементов  $N$ , число блоков  $M$ , число эле-

ментов в каждом блоке  $k$  и число  $m$  блоков, в которое входит каждый элемент. Блок-схема задается набором блоков, в которых элементы размещены согласно следующему условию:  $N$  элементов размещены по  $M$  блокам так, что каждый блок содержит *точно*  $k$  различных элементов, каждый элемент появляется *точно* в  $m$  различных блоках и каждая пара элементов входит *точно* в один блок. Теория рассматривает случай  $k < N$ , когда размер блока меньше числа элементов, и  $N \leq M$ , когда число элементов не больше числа блоков.

Необходимые условия существования УБС задаются выражениями [12]:

$$\begin{aligned} Nm &= kM, \\ m(k - 1) &= N - 1, \\ k &\leq m, k < N. \end{aligned} \quad (1)$$

Минимальное число блоков и максимальное число элементов обеспечиваются в случае симметричных блок-схем, когда  $m = k$  и  $M = N$ . Они имеют всего два параметра: фактор вхождения  $m$  и число элементов или блоков  $N = m(m - 1) + 1$ . Достаточное условие существования симметричной блок-схемы заключается в выполнении соотношения  $m = p^r + 1$ , где  $p$  — простое и  $r$  — натуральное числа.

В табл. 2 и 3 представлены симметричные блок-схемы при значениях  $m = 2$  и  $m = 3$  соответственно [12, 15, 16].

Теперь заменим в определении УБС блок на малую сеть, элемент — на абонента сети (процессор), вхождение элемента в блок — на дуплексное подсоединение абонента к малой сети и, наконец, блок-схему — на большую сеть. Тогда мы получим следующее определение большой сети — это сеть, к которой подсоединено  $N$  абонентов и которая состоит из  $M$  малых сетей, к каждой из которых подсоединено *точно*  $k$  разных абонентов, каждый абонент входит *точно* в  $m$  малых сетей и

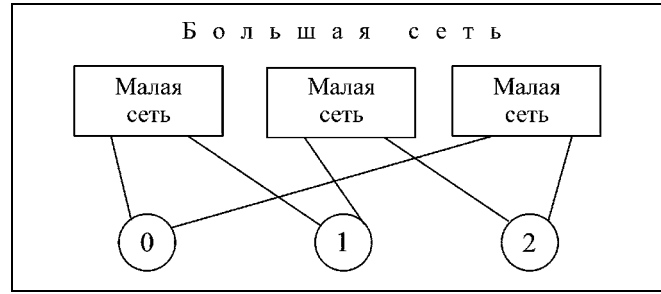


Рис. 2. Однородное распараллеливание большой сети или инвариантное расширение малой сети

каждый абонент соединяется с любым другим абонентом *точно* через одну малую сеть.

Из этого определения следует, что в большой сети любые два абонента соединены друг с другом *точно* также как в малой сети, что означает сохранение в большой сети таких базовых функциональных свойств малой сети, как способ маршрутизации и задержка передачи данных при одновременном увеличении числа объединяемых абонентов и, значит, пропускной способности.

Рассмотренный метод распараллеливания большой сети на ряд малых сетей можно также трактовать и как однородное расширение малой сети. Он имеет своим инвариантом способ маршрутизации и задержку передачи.

На рис. 2 показан способ распараллеливания большой сети или расширения малой сети согласно УБС из табл. 2.

Большая сеть, являющаяся *точным* аналогом блок-схемы, названа нами простейшей сетью (ПС). К сожалению, ПС имеют дискретные и небольшие размеры [12, 16, 17]. Их параметры представлены в табл. 4. Поэтому необходимо уметь представлять сеть любого

Таблица 2

Блок-схема с  $m = k = 2, M = N = 3$

| Блоки | Элементы |   |
|-------|----------|---|
| 0     | 0        | 2 |
| 1     | 1        | 0 |
| 2     | 2        | 1 |

Таблица 3

Блок-схема с  $m = k = 3, M = N = 7$

| Блоки | Элементы |   |   |
|-------|----------|---|---|
| 0     | 0        | 6 | 4 |
| 1     | 1        | 0 | 5 |
| 2     | 2        | 1 | 6 |
| 3     | 3        | 2 | 0 |
| 4     | 4        | 3 | 1 |
| 5     | 5        | 4 | 2 |
| 6     | 6        | 5 | 3 |

Таблица 4

Число абонентов  $N$  в различных блок-схемах или ПС

| M  | K  |    |    |    |    |    |    |    |    |     |     |  |
|----|----|----|----|----|----|----|----|----|----|-----|-----|--|
|    | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11  | 12  |  |
| 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11  | 12  |  |
| 2  | 3  | —  | —  | —  | —  | —  | —  | —  | —  | —   | —   |  |
| 3  | 4  | 7  | —  | —  | —  | —  | —  | —  | —  | —   | —   |  |
| 4  | 5  | 9  | 13 | —  | —  | —  | —  | —  | —  | —   | —   |  |
| 5  | 6  | —  | 16 | 21 | —  | —  | —  | —  | —  | —   | —   |  |
| 6  | 7  | 13 | —  | 25 | 31 | —  | —  | —  | —  | —   | —   |  |
| 7  | 8  | 15 | —  | —  | —  | —  | —  | —  | —  | —   | —   |  |
| 8  | 9  | —  | 25 | —  | —  | 49 | 57 | —  | —  | —   | —   |  |
| 9  | 10 | 19 | 28 | —  | —  | —  | 64 | 73 | —  | —   | —   |  |
| 10 | 11 | 21 | —  | 41 | —  | —  | —  | 81 | 91 | —   | —   |  |
| 11 | 12 | —  | —  | 45 | —  | —  | —  | —  | —  | —   | —   |  |
| 12 | 13 | 25 | 37 | —  | —  | —  | —  | —  | —  | 121 | 133 |  |
| 13 | 14 | 27 | 40 | —  | —  | —  | —  | —  | —  | —   | —   |  |



размера через ПС. Наличие такого представления открывает возможность применения теории блок-схем для любых системных сетей.

### 3. РАЗЛОЖЕНИЕ НА ПРОСТЕЙШИЕ СЕТИ

Как видно из табл. 4 и формулы (1), простейшие сети, изоморфные УБС, имеют число абонентов  $N$  в диапазоне  $2 \leq N \leq m(m-1) + 1$  и достигают максимума по  $N$  в симметричных ПС при  $m = k$ . В системных сетях значение  $m$  задает степень разветвления входного/выходных каналов абонента (процессора) и, поэтому, не может быть очень большим (например, не должно выходить за рамки табл. 4). Для преодоления отмеченного ограничения на величину  $N$  в работах [15, 16] предложено использовать ПС в качестве строительных блоков, из которых составляется большая сеть общего вида заданного размера.

Пусть мы имеем  $M$  одинаковых малых сетей, каждая из которых объединяет  $K$  абонентов ( $K > k$ , где  $k$  — параметр УБС и ПС). Необходимо построить большую сеть, состоящую из  $M$  малых сетей, объединяющую  $R > K$  абонентов и имеющую функциональные свойства малой сети.

Для этого разобьем каждую малую сеть на  $\lfloor K/k \rfloor$  сверхмалых сетей и перенумеруем их в каждой малой сети в диапазоне  $0 \leq j < \lfloor K/k \rfloor$ . Объединим все  $j$ -е сверхмалые сети в  $j$ -ю ПС, а 0-я ПС должна иметь стандартную структуру. К  $j$ -й ПС подсоединим абонентов с номерами от  $jN$  до  $N(j+1) - 1$  так, чтобы они на одинаковых позициях имели номера на  $N$  больше, чем номера в  $(j-1)$ -й ПС.

Таким образом, мы получим большую сеть, состоящую из  $\lceil K/k \rceil$  ПС, объединяющую  $R$  абонентов, где  $R$  находится в диапазоне:  $N \lfloor K/k \rfloor \leq R \leq N \lceil K/k \rceil$ . В табл. 5 показан этот процесс для  $m = k = 3$ ,  $M = N = 7$  и  $K = 16$ . Различные малые сети размещены по строкам таблицы, а различные ПС — в прямоугольных областях, выделенных различными шрифтами. Если  $\lceil K/k \rceil \neq \lfloor K/k \rfloor$ , то последняя  $\lfloor K/k \rfloor$ -я ПС — усеченная. Два разных способа построения усеченных ПС рассмотрены в работах [15, 16]. В результате проведенного однородного параллельного расширения системной сети число абонентов в ней увеличилось от  $K = 16$  до  $R = 36$  (см. табл. 5).

Полученная в результате построения большая сеть общего вида является неуравновешенной плоской сетью, набранной из малых сетей. Для нее справедлива следующая теорема.

**Таблица 1.** Большая сеть общего вида представляет собой плоскую сеть малых сетей, в которой между любыми двумя абонентами существует хотя бы один путь, проходящий через некоторую исходную сеть. Таких путей может быть несколько (в отличие от расширенных ПС). Все они проходят через разные исходные сети.

**Доказательство.** Каждую ПС составляют абоненты, номера которых образуют кольцо вычетов по  $\text{mod } N$ . Между абонентами любой ПС, номера которых имеют разные вычеты, существуют пути только через те исходные сети, что и в 0-й ПС. Между абонентами разных ПС, номера которых имеют разные вычеты, существуют пути только через те исходные сети, что и в отдельной ПС. Между абонентами разных ПС, номера которых имеют одинаковые вычеты, существует  $m$  разных путей, которые проходят через разные исходные сети. Это следует из того факта, что указанные абоненты подсоединены к одним и тем же исходным сетям. ♦

По проведенному построению в большой сети каждый абонент соединяется с любым другим абонентом не менее чем через одну малую сеть. Это обеспечивает сохранение (не ухудшение) в большой сети функциональных свойств малой сети.

### 4. МЕТОД ИНВАРИАНТНОГО РАСШИРЕНИЯ СИСТЕМНЫХ СЕТЕЙ

Данный метод можно применять при построении системных сетей как сверху вниз, так и снизу вверх. В первом случае его можно называть методом однородного распараллеливания большой сети, а во втором — методом однородного расширения малой сети. Во втором случае можно указать инвариант расширения — способ маршрутизации и его характеристики (конфликтность, пропускную способность, задержки передачи и т. п.). Здесь используется, в основном, второй подход.

Пусть имеется некоторая системная сеть  $S_K$ , объединяющая  $K$  абонентов (процессоров). Необходимо построить расширенную системную сеть  $S_R$ , объединяющую  $R > K$  абонентов.

Таблица 5

Большая сеть с  $m = k = 3$ ,  $M = N = 7$  и  $K = 16$ ,  $R = 36$

| Номер малой сети | Большая сеть |          |          |     |    |    |           |           |           |     |    |    |           |           |           |     |
|------------------|--------------|----------|----------|-----|----|----|-----------|-----------|-----------|-----|----|----|-----------|-----------|-----------|-----|
|                  | ПС           |          |          |     |    |    |           |           |           |     |    |    |           |           |           |     |
|                  | 0-я          |          |          | 1-я |    |    | 2-я       |           |           | 3-я |    |    | 4-я       |           |           | 5-я |
| 0                | <b>0</b>     | <b>6</b> | <b>4</b> | 7   | 12 | 8  | <b>14</b> | <b>20</b> | <b>18</b> | 21  | 27 | 25 | <b>28</b> | <b>34</b> | <b>32</b> | 35  |
| 1                | <b>1</b>     | <b>0</b> | <b>5</b> | 8   | 13 | 9  | <b>15</b> | <b>14</b> | <b>19</b> | 22  | 21 | 26 | <b>29</b> | <b>28</b> | <b>33</b> | 35  |
| 2                | <b>2</b>     | <b>1</b> | <b>6</b> | 9   | 7  | 10 | <b>16</b> | <b>15</b> | <b>20</b> | 23  | 22 | 27 | <b>30</b> | <b>29</b> | <b>34</b> | —   |
| 3                | <b>3</b>     | <b>2</b> | <b>0</b> | 10  | 8  | 11 | <b>17</b> | <b>16</b> | <b>14</b> | 24  | 23 | 21 | <b>31</b> | <b>30</b> | <b>28</b> | 35  |
| 4                | <b>4</b>     | <b>3</b> | <b>1</b> | 11  | 9  | 12 | <b>18</b> | <b>17</b> | <b>15</b> | 25  | 24 | 22 | <b>32</b> | <b>31</b> | <b>29</b> | —   |
| 5                | <b>5</b>     | <b>4</b> | <b>2</b> | 12  | 10 | 13 | <b>19</b> | <b>18</b> | <b>16</b> | 26  | 25 | 23 | <b>33</b> | <b>32</b> | <b>30</b> | —   |
| 6                | <b>6</b>     | <b>5</b> | <b>3</b> | 13  | 11 | 7  | <b>20</b> | <b>19</b> | <b>17</b> | 27  | 26 | 24 | <b>34</b> | <b>33</b> | <b>31</b> | —   |

Пусть дуплексное подсоединение каждого абонента к сети  $S_R$  осуществляется через разветвитель входа/выхода на  $m$  каналов  $B_m$ . Роль такого разветвителя могут выполнять  $m$  сетевых карт на одной шине ввода/вывода (PCI для персональных компьютеров), или  $m$ -портовая сетевая карта, или, в общем случае, — отдельные схемы, объединенные в специальную СБИС с перестраиваемой степенью ветвления  $m$ . Разветвитель  $B_m$  должен выполнять две функции. На выходе из абонента — функцию направления кадра в необходимую малую сеть. При этом разветвитель может сам выбирать малую сеть (маршрутная адресация), а может передавать эту функцию малым сетям (сетевая адресация). На входе в абонент разветвитель должен выполнять функцию множественного доступа в общий канал, т. е. выбирать один из многих каналов для передачи кадра. Схемы, выполняющие такую функцию, могут иметь минимальную глубину (например, два логических элемента [18]).

Построим сеть  $S_R$  как большую сеть из  $M$  параллельно включенных малых сетей  $S_K$  с разбиением последних на  $\lceil K/k \rceil$  простейших сетей ширины  $k$ , объединяющих согласно формуле (1)  $N = m(k-1) + 1$  абонентов. Разбиение на простейшие сети, описанное в § 3, задает и схему подсоединения абонентов к сетям  $S_K$ . Теория блок-схем утверждает, что максимум  $N = M$  и минимум  $M$  достигается при  $k = m$ . Выбирая это значение  $k$ , мы получаем, что расширенная сеть  $S_R$  объединяет  $R$  абонентов, где  $\lfloor K/m \rfloor (m(m-1) + 1) \leq R \leq \lceil K/m \rceil (m(m-1) + 1)$ . Если  $m$  является делителем  $K$ , то коэффициент расширения  $\alpha$  системной сети задается выражением  $\alpha = R/K = (m(m-1) + 1)/m$ . При расширении сети происходит ее усложнение, которое можно задавать коэффициентом  $\beta = m(m-1) + 1$ . В табл. 6 представлены значения этих коэффициентов для ряда малых значений  $m = k$ .

Из построения видно, что задержки в расширенной сети  $S_R$  больше задержек в исходной сети  $S_K$  на размер задержек в разветвителе каналов  $B_m$ . Поэтому утверждение об инвариантности по задержкам справедливо лишь с указанной точностью. Однако эта точность может быть достаточно велика, так как схема разветвителя является чисто комбинационной и может быть сделана очень малой глубины, задержки на которой будут много меньше, чем в исходной сети.

Рассматриваемый метод допускает и итеративное многошаговое применение, при котором в качестве исходной сети берется расширенная на предыдущем шаге сеть. Однако применение итеративного расширения ограничивается более быстрым ростом числа исходных сетей на каждом шаге по сравнению с ростом числа абонентов (см. табл. 6). Поэтому его можно рекомендовать только для тех случаев, когда необходимое расширение не может быть получено за один шаг.

В работах [15, 16] рассмотрено применение рассматриваемого метода для таких системных сетей, как коммутируемые локальные сети [15] и однокаскадные коммутаторы [16]. Рассмотрим применение данного метода для некоммутируемых мультиколец (см. § 1).

Рассмотрим исходную сеть, состоящую из пары встречных колец с  $K$  узлами (процессорами). При передаче кадров в кольцо с кратчайшим маршрутом эта сеть имеет емкость  $c = 8$ . Для расширения исходной сети будем использовать в каждом процессоре разветвитель каналов  $B_m$ . Пусть  $m$  является делителем  $K$ . Расширенная сеть содержит  $M = N = m(m-1) + 1$  исходных сетей, объединяет  $R = KN/m$  процессоров и имеет емкость  $C = cM = 8Rm/K$ . Последовательность подсоединения процессоров к парам колец задается процедурой разбиения  $M$  исходных сетей на простейшие сети (см. § 3).

Здесь построено другое мультикольцо, чем описанное в § 1, — с большим числом колец меньшего размера и с большей емкостью. В нем все кольца имеют переменный шаг между узлами и одинаковую структуру колец. Однако это мультикольцо удалось исследовать сразу аналитически. В частности, при  $K = 16$  и  $m = 2$  имеем  $M = 3$  (число колец 6),  $R = 24$  и  $C = R$ ; при  $K = 16$  и  $m = 4$  —  $M = 13$  (число колец 26);  $R = 52$  и  $C = 2R$ ; при  $K = 15$  и  $m = 5$  —  $M = 21$  (число колец 42),  $R = 63$  и  $C = 8R/3$ .

Для дальнейшего увеличения числа процессоров и емкости сети в качестве исходной сети можно брать мультикольцо из табл. 1. В этом случае  $c = K$  и  $C = Rm$ . В частности, если исходная сеть содержит  $K = 64$  процессора и состоит из 8 колец, то при  $m = 2$  —  $M = 3$  (24 кольца),  $R = 96$  и  $C = 2R$ ; а при  $m = 4$  —  $M = 13$  (104 кольца),  $R = 208$  и  $C = 4R$ .

## 5. ВОЗМОЖНОСТИ ПРАКТИЧЕСКОГО ПРИМЕНЕНИЯ

Рассмотренный метод может иметь очень широкое применение — буквально на каждом уровне иерархии системной сети. Рассмотрим два крайних случая — уровень коммутационных СБИС и уровень всей сети.

На рис. 3 показана схема построения новой более широкой СБИС из трех старых СБИС согласно УБС из табл. 2. Здесь роль малой сети выполняет старая СБИС, большой сети — новая СБИС, а абонентов — контакты  $B/V$ . В схеме используются разветвители каналов  $B_2$  и показаны каждые два контакта старой СБИС и каждые три контакта новой СБИС. Новая СБИС будет иметь в 1,5 раза большее число контактов, чем старая СБИС. Если использовать 7 старых СБИС, то число контактов увеличится в 2,3 раза, и т. д. по табл. 6.

Аналогично для полной системной сети. Возьмем, например, самую мощную параллельную вычислительную систему Blue Gene фирмы IBM, содержащую 131

Таблица 6

Коэффициенты расширения и усложнения системной сети

| $m$      | 2   | 3    | 4    | 5   | 6    | 8     | 9    | 10  | 12    | 14    | 17    | 18    |
|----------|-----|------|------|-----|------|-------|------|-----|-------|-------|-------|-------|
| $\beta$  | 3   | 7    | 13   | 21  | 31   | 55    | 83   | 91  | 133   | 183   | 273   | 307   |
| $\alpha$ | 1,5 | 2,33 | 3,25 | 4,2 | 5,17 | 7,125 | 8,11 | 9,1 | 11,08 | 13,07 | 16,06 | 17,06 |

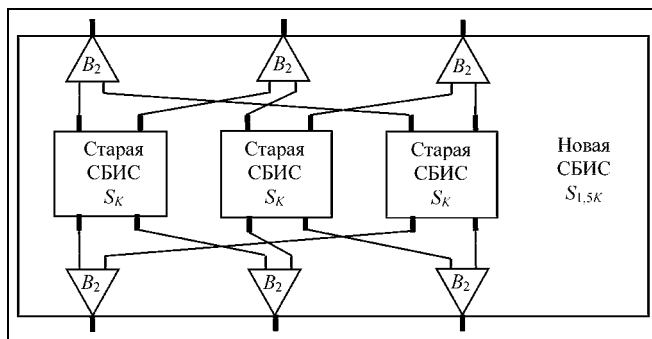


Рис. 3. Инвариантное увеличение числа контактов СБИС-коммутатора

тыс. процессоров. Ее системная сеть — это трехмерный тор (куб со сшитыми краями). Заменяем однопортовую сетевую карту в каждом процессоре на двухпортовую. Тогда чисто механически по схеме (см. рис. 2), используя три системных сети, можно увеличить число процессоров до 195 тыс. Если использовать трехпортовую сетевую карту и семь системных сетей, то можно увеличить число процессоров до 306 тыс., и т. д. по табл. 6. Предел здесь ставит соотношение сложности процессорной и сетевой подсистем.

#### ЗАКЛЮЧЕНИЕ

Простота и универсальность рассмотренного метода расширения системных сетей и наличие разработанной математической теории делают данный метод очень привлекательным для дальнейших исследований. Наиболее привлекательным направлением исследований может стать разработка метода построения отказоустойчивых системных сетей [19].

Дело в том, что общее определение блок-схемы [12] содержит еще один параметр  $\sigma \geq 1$  — число блоков, в которые входит каждая пара элементов. При этом блок-схема задается набором блоков, в которых каждая пара элементов входит *точно* в  $\sigma$  блоков. Тогда большой сетью является та сеть, к которой подсоединено  $N$  абонентов и которая состоит из  $M$  малых сетей, к каждой из которых подсоединено *точно*  $k$  разных абонентов, каждый абонент входит *точно* в  $t$  малых сетей и каждый абонент соединяется с любым другим абонентом *не менее* чем через  $\sigma$  малых сетей. При таком определении большая сеть является устойчивой к отказу ( $\sigma - 1$ ) малых сетей. Поэтому теория блок-схем может служить математической основой для построения отказоустойчивых системных сетей.

#### ЛИТЕРАТУРА

1. Медведев И. Л., Прангшвили И. В., Чудин А. А. Многопроцессорные вычислительные системы с перестраиваемой структурой. — М.: Ин-т пробл. управления. — 1975. — 87 с.

2. Прангшвили И. В., Подлазов В. С., Стецюра Г. Г. Локальные микропроцессорные вычислительные сети. — М.: Наука, 1984. — 176 с.
3. Алленов А. В., Подлазов В. С., Стецюра Г. Г. Пропускная способность набора кольцевых каналов. I. Класс наборов колец. Наборы с простыми узлами // Автоматика и телемеханика. — 1996. — № 3. — С. 135—144.
4. Алленов А. В., Подлазов В. С. Пропускная способность набора кольцевых каналов II. Кольцевые коммутаторы // Там же. — № 4. — С. 162—172.
5. Подлазов В. С., Подлазова А. В. Обеспечение наращиваемости отказоустойчивых многопроцессорных систем с общей памятью с использованием многокольцевых некоммутируемых сетей связи с неоднородными узлами // Тр. Ин-та пробл. упр. РАН. — 2002. — Т. XVIII. — С. 164—181.
6. Подлазов В. С. Наращиваемые многокольцевые некоммутируемые сети связи для многопроцессорных вычислительных систем // Проблемы управления. — 2006. — № 2. — С. 50—57.
7. Подлазов В. С. Бесконфликтная статическая децентрализованная маршрутизация для кольцевых коммутаторов и гиперкубов // Автоматика и телемеханика. — 1999. — № 4. — С. 79—89.
8. Подлазов В. С. Неблокируемые кольцевые коммутаторы: Резервирование и быстродействие // Там же. — № 10. — С. 153—163.
9. Подлазов В. С. Условия неблокируемости мультикольцевых коммутаторов и обобщенных гиперкубов на произвольных перестановках. I. Межузловая коммутация. Мультикольца // Там же. — 2001. — № 8. — С. 118—126.
10. Подлазов В. С. Условия неблокируемости мультикольцевых коммутаторов и обобщенных гиперкубов на произвольных перестановках. II. Обобщенные гиперкубы. Внутриузловая коммутация // Там же. — № 9. — С. 114—124.
11. Подлазов В. С. Неблокируемость произвольной перестановки на коммутаторе со структурой многомерной решетки // Там же. — 2002. — № 8. — С. 178—188.
12. Холл М. Комбинаторика. — М.: Мир, 1970. — 424 с.
13. Rzymianowicz L. Designing efficient network interfaces for system area networks < [http://bibserv7.bib.uni-mannheim.de/madoc/volltexte/2002/54/pdf/54\\_1.pdf](http://bibserv7.bib.uni-mannheim.de/madoc/volltexte/2002/54/pdf/54_1.pdf)>.
14. Уваров С. И., Чернышев М. А. О предельном быстродействии кольцевых устройств обмена информацией для параллельных вычислительных систем // Автоматика и телемеханика. — 1986. — № 10. — С. 151—156.
15. Подлазов В. С., Стецюра Г. Г. Регулярные плоские сети для суперкомпьютеров // Проблемы управления. — 2006. — № 1. — С. 26—31.
16. Подлазов В. С., Соколов В. В. Однокаскадные коммутаторы большой размерности для многопроцессорных и многомашиных вычислительных систем // Там же. — № 6. — С. 19—24.
17. Finn A. M., Decker R. O. A network architecture for radar signal processing // AIAA/IEEE 8-th digital avionic systems conference / San Jose. Calif. Oct., 1988. — P. 614—621.
18. Prangishvili I. V., Stetsyura G. G., Podlazov V. S. Decentralized control of processor interactions in concentrated and distributed multimicroprocessor system // Microprocessing and Microprogramming. — 1987. — N 7. — P. 220—228.
19. Каравай М. Ф. Инвариантно-групповой подход к исследованию  $k$ -устойчивых структур // Автоматика и телемеханика. — 2000. — № 1. — С. 144—156.

☎ (495) 334-78-31

e-mail: podlazov@ipu.ru

Статья представлена к публикации членом редколлегии В. Д. Малюгиным. □