УДК 004.724.2 + 004.272.43 DOI: http://doi.org/10.25728/pu.2020.3.7

# ОТКАЗОУСТОЙЧИВЫЙ НЕБЛОКИРУЕМЫЙ ТРЕХМЕРНЫЙ РАЗРЕЖЕННЫЙ ГИПЕРКУБ

В.С. Подлазов

**Аннотация.** Предложена системная сеть в виде неблокируемого отказоустойчивого трехмерного обобщенного *p*-ичного гиперкуба с одним процессором в каждом узле гиперкуба. В любой неблокируемой сети данные между абонентами передаются с наименьшими задержками по прямым каналам без промежуточной буферизации. Отмечено, что сети с топологией обобщенного гиперкуба имеют наименьшие длины прямых каналов и наименьшие задержки передачи. Разработана структура этого гиперкуба на базе сетей с топологией квазиполного графа, которая позволяет разменивать число абонентов на число разных прямых каналов между любыми абонентами и задавать канальную и узловую отказоустойчивость сети. Даны параметры квазиполных графов, существующих при любой *p*-ичности гиперкуба. В результате предложена структура сети в виде разреженного *p*-ичного гиперкуба с числом узлов несколько меньшим, чем в обычном *p*-ичном гиперкубе. Разреженный гиперкуб разработан как отказоустойчивая системная сеть для однокристального процессора-ускорителя с несколькими сотнями ядер. Разработаны процедура и алгоритм прокладки бесконфликтных прямых каналов посредством динамической локальной самомаршрутизации пакетов, при котором узлы не взаимодействуют друг с другом и используют только расширенную маршрутную информацию из пакетов.

**Ключевые слова:** системная сеть, квазиполный граф, обобщенный гиперкуб, канальная отказоустойчивость, коммутационные свойства, неблокируемая сеть, бесконфликтные прямые каналы, локальная динамическая самомаршрутизация.

## ВВЕДЕНИЕ

Данная статья продолжает серию работ автора по созданию самомаршрутизируемых неблокируемых системных сетей для многопроцессорных вычислительных систем [1—3]. Эти сети обладают наибольшим быстродействием, так как передача данных в них ведется по прямым каналам без промежуточной буферизации.

Исходно эти сети имели топологию квазиполных графа или орграфа [1, 2] и обладали квадратичной зависимостью числа узлов от их степени. Самомаршрутизация в этих сетях статическая, при которой любой источник самостоятельно прокладывает весь бесконфликтный путь к приемнику для произвольной перестановки пакетов данных между узлами. Главный недостаток этих сетей — малое число узлов в них. Топологию квазиполного орграфа имели сети в виде 2-мерного обобщенного гиперкуба и 2-мерного полного мультикольца [2].

Обобщенный *r*-мерный *p*-ичный гиперкуб [4] со степенью узлов $m = rp$ и числом узлов $N = p^r$ применяется как системная сеть в некоторых современных многопроцессорных вычислительных системах [5]. В работе [3] на базе 2-мерного гиперкуба был построен неблокируемый 3-мерный *p*-ичный гиперкуб, для которого был предложен алгоритм динамической локальной самомаршрутизации. Он позволяет параллельно прокладывать бесконфликтные пути при произвольной перестановке пакетов данных между узлами на основе только локальной информации в промежуточных узлах без взаимодействия между ними.

Полное *r*-мерное *p*-ичное мультикольцо также имеет $N = p^r$ узлов со степенью $m = rp$. В работе [3] на базе 2-мерного мультикольца было построено неблокируемое 3-мерное *p*-ичное мультикольцо, алгоритм динамической локальной самомаршрутизации в котором имел те же свойства, что и в гиперкубе. Эта работа была ориентирована на проектирование сетей для однокристальных процессоров-ускорителей с кольцевыми каналами типа *Cell*, *Xion Phy* или ПС2100, в которых необходимо иметь сотни и тысячи ядер.

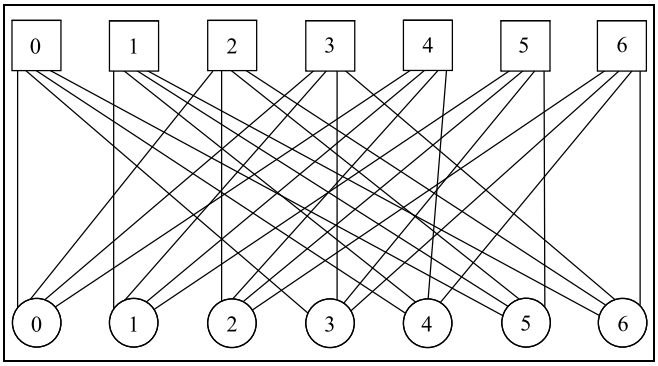


**Рис. 1. Квазиполный граф с $N = 7$, $p = 4$ и $\sigma = 2$:** квадраты — коммутаторы

К сожалению, сети с топологией квазиполных орграфов не имеют изначальных (врожденных) свойств обеспечения их отказоустойчивости. Тогда как сети с топологией квазиполного графа таким свойствами обладают — в них возможно создание конфигураций с несколькими (двумя, тремя и более) разными путями между любыми двумя узлами. Поэтому в работе [3] была поставлена задача создания на базе $p$-ичного квазиполного графа отказоустойчивой 3-мерной неблокируемой сети с числом узлов $R = O(p^3)$ и с наличием $\sigma$ разных прямых путей между узлами.

В настоящей статье предлагается структура расширенного по каналам 3-мерного $p$-ичного гиперкуба, который образует неблокируемую системную сеть с канальной отказоустойчивостью заданного порядка. Для этой сети предложен алгоритм динамической самомаршрутизации пакетов, который прокладывает бесконфликтные пути на основе расширенной маршрутной информации локально по узлам сети.

*Структура работы.* Во Введении обосновывается актуальность построения отказоустойчивых неблокируемых гиперкубов, так как ранее построенные неблокируемые гиперкубы и мультикольца не обладали канальной отказоустойчивостью. В § 1 вводятся такие основные понятия, как *квазиполный* граф, *плоская сеть* и бесконфлитктная *самомаршрутизация.* В § 2 дается новое основное понятие данной статьи, а именно *3-мерный разреженный гиперкуб из плоских сетей.* Даются структуры гиперкуба и составляющих его плоских сетей. В § 3 объясняется, как в построенном гиперкубе осуществляется *отказоустойчивая* самомаршрутизация по разным путям через плоские сети. В § 4 объясняется, как в такой топологии сети организуется *неблокируемая* самомаршрутизация. Обоснование неблокируемости требует введения дополнительных параллельных каналов в структуру гиперкуба и рассмотрения методов разрешения на их основе двух типов возникающих конфликтов. Наконец, в § 5 описывается алгоритм неблокируемой самомаршрутизации, который можно непосредственно реализовывать в коммутаторах сети. Алгоритм использует передаваемую в пакетах маршрутную информацию *динамически* локально по узлам, т. е. в процессе прокладки бесконфликтных путей. В Заключении отражены основные результаты статьи.

## 1. НЕБЛОКИРУЕМАЯ ОТКАЗОУСТОЙЧИВАЯ СЕТЬ С ТОПОЛОГИЕЙ КВАЗИПОЛНОГО ГРАФА

*Квазиполный граф* представляет собой минимальный двудольный граф, одну долю которого составляют узлы-абоненты, а другую — узлы-коммутаторы (рис. 1). Степень всех узлов в каждой доле одинакова и равна $p$. Число вершин в каждой доле $N$, перенумерованных от 0 до $N - 1$, задается равенством $N = p(p - 1)/\sigma + 1$ [6], где $\sigma$ задает число разных путей между любыми двумя абонентами. Каждый такой путь имеет длину в 2 ребра через один и только один коммутатор. Каждое ребро представляет собой две встречных дуги — от абонента к коммутатору и обратно.

Схемы соединений между коммутаторами и абонентами задаются, например, в табл. 1. В каждой строке первая ячейка задает номер коммутатора, а остальные ячейки — номера абонентов, связанных с данным коммутатором ребром в виде дуплексного канала (двух встречных симплексных каналов).

Если объединить одноименные коммутатор и абонента квазиполного графа в одном сетевом узле, то мы получим *плоскую сеть* ПС($N$, $p$, $\sigma$). Она является неблокируемой сетью с бесконфликтной статической самомаршрутизацией и обладает ($\sigma - 1$)-канальной отказоустойчивостью при $\sigma > 1$ (рис. 2).

Самомаршрутизация в ПС($N$, $p$, $\sigma$) осуществляется по схеме соединений (см. табл. 1) с помощью номеров узлов абонента-источника и абонента-приемника. В таблице находятся номера комму-

*Таблица 1*

**Схемы соединений в двух квазиполных графах**

| Коммутаторы | $N = 7$ $p = 3$, $\sigma = 1$ | | | $N = 7$ $p = 4$, $\sigma = 2$ | | | |
|---|---|---|---|---|---|---|---|
| | Абоненты | | | Абоненты | | | |
| 0 | 0 | 6 | 4 | 0 | 5 | 4 | 3 |
| 1 | 1 | 0 | 5 | 1 | 6 | 5 | 4 |
| 2 | 2 | 1 | 6 | 2 | 0 | 6 | 5 |
| 3 | 3 | 2 | 0 | 3 | 1 | 0 | 6 |
| 4 | 4 | 3 | 1 | 4 | 2 | 1 | 0 |
| 5 | 5 | 4 | 2 | 5 | 3 | 2 | 1 |
| 6 | 6 | 5 | 3 | 6 | 4 | 3 | 2 |

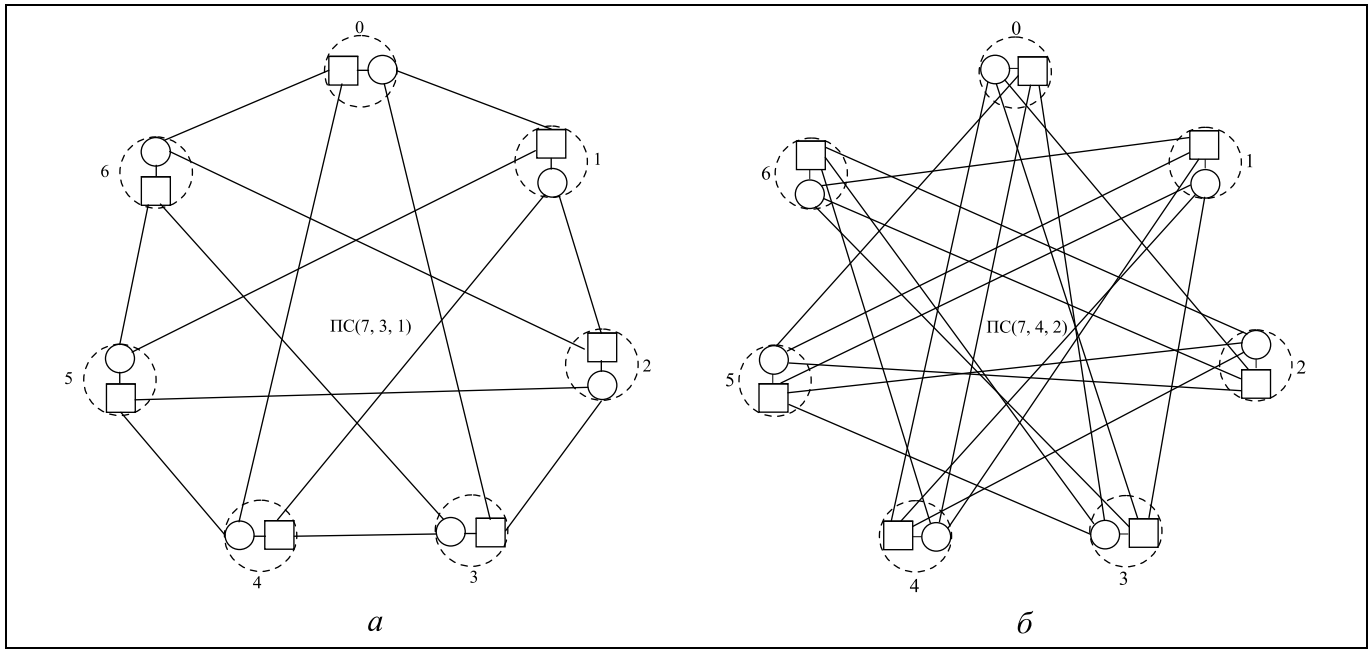


**Рис. 2. Неблокируемые сети:** *а* — ПС(7, 3, 1); *б* — ПС(7, 4, 2)

таторов, связанных дуплексными каналами с абонентами-приемниками. Среди этих коммутаторов находится тот, который связан таким же каналом с абонентом-источником. По свойствам квазиполных графов таких коммутаторов σ.

На самом деле, каждая таблица соединений представляет собой одну из неполных уравновешенных симметричных блок-схем $BD(N, p, \sigma)$, исследуемых в комбинаторике [7]. Они существуют и, тем более, построены не для всех значений параметров $p$ и σ. В табл. 2 приведены значения $N$ для блок-схем при малых значениях этих параметров. Пустые клетки отмечают блок-схемы, которые не существуют по определению. Прочерки в клетках отмечают блок-схемы, которые не могут существовать по теории, а перечеркнутые значения отмечают блок-схемы, которые еще не построены.

*Таблица 2*

**Число элементов в блок-схемах $BD(N, p, \sigma)$**

| σ | *p* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | 7 | 11 | 21 | 31 | — | 57 | 73 | 91 | ~~111~~ | 133 |
| 2 | 4 | 7 | 11 | 16 | — | — | 37 | — | ~~56~~ | — |
| 3 | 3 | 5 | | 11 | 15 | | 25 | 31 | ~~45~~ | |

*Таблица 3*

**Число элементов $N$ в блок-схемах $BD(N, p, \sigma)$ и $N^*$ в 1-расширенных блок-схемах $BD(N^*, p, \sigma|\sigma + 1)$**

| σ | *p* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | 7 | 11 | 21 | 31 | 39 | 57 | 73 | 91 | 95 | 133 |
| 2 | 4 | 7 | 11 | 15 | 21 | 27 | 37 | 42 | 51 | 63 |
| 3 | 3 | 5 | 7 | 11 | 15 | 19 | 23 | 29 | 36 | 43 |

Необходимость построения отказоустойчивых сетей ПС($N$, $p$, σ) требует некоторого эффективного заполнения пустых клеток в табл. 2. Для этого в работе [8] были предложены и построены 1-расширенные блок-схемы $BD(N^*, p, \sigma|\sigma + 1)$, задающие таблицы соединений для *1-расширенных плоских сетей* ПС($N^*$, $p$, σ|σ + 1), в которых малая часть абонентов связаны σ + 1 разными путями, а остальные — точно σ разными путями. Значения $N$ и $N^*$ числа узлов в упомянутых блок-схемах приведены в табл. 3, где последние выделены подчеркиванием.

## 2. ТРЕХМЕРНЫЙ РАЗРЕЖЕННЫЙ ГИПЕРКУБ ИЗ ПЛОСКИХ СЕТЕЙ

Пусть имеется сеть из $N$ узлов с номерами $(x_1, x_2, ..., x_N)$, где $x_i \in [0, Np - 1]$. *Следом* $t(x_i)$ узла $x_i$ считается остаток от деления его номера на $N$. Такая сеть считается *плоской*, если в ней любые два узла со следами $t(x_i)$ и $t(x_j)$ объединяются такими же ребрами (учетом абонентов и коммутаторов), как узлы с теми же следами в сети ПС($N$, $p$, σ).

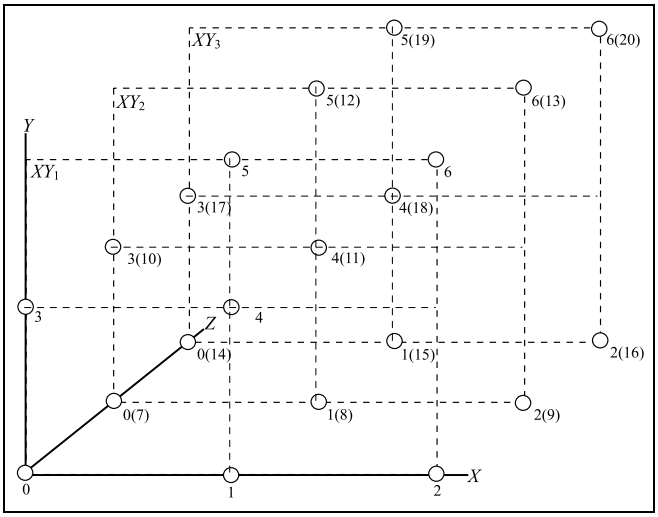


**Рис. 3. Разреженный троичный гиперкуб с плоскими сетями на плоскостях $XY_i$**

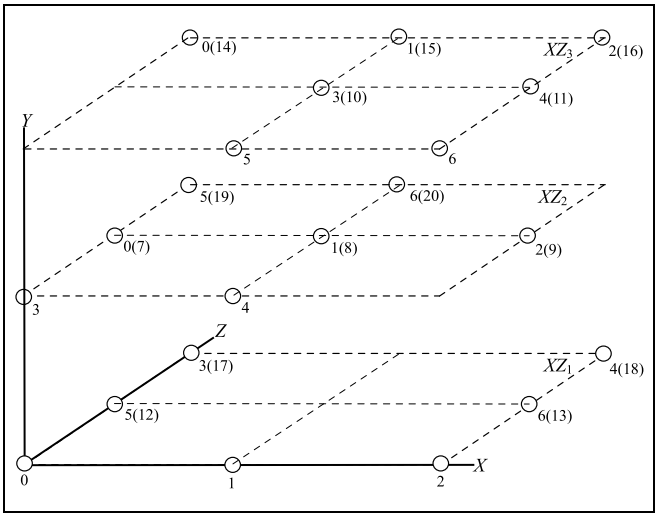


**Рис. 4. Разреженный троичный гиперкуб из плоских сетей на плоскостях $XY_i$ и $XZ_j$**

Иначе говоря, плоскими мы считаем сети, изоморфные сети ПС($N$, $p$, σ).

*Шаблоном* $T$ плоской сети с номерами узлов ($x_1$, $x_2$, ..., $x_N$) мы называет набор (множество) их следов $T = \{t(x_1), t(x_2), ..., t(x_N)\}$.

Построим *трехмерный p-ичный гиперкуб* [3], каждая плоскость $XY$ и $XZ$ которого содержит плоские сети из $p$ узлов, с единым шаблоном $T$ сетей ПС($N$, $p$, σ) или ПС($N^*$, $p$, σ|σ + 1). Не снижая общности, построение будем вести на примере ПС(7, 3, 1) (см. табл. 1), в которой номера узлов и их следы совпадают. В дальнейшем всегда будем применять обозначение ПС($N$, $p$, σ), даже если используется ПС($N^*$, $p$, σ|σ + 1).

Будем размещать на вертикальных плоскостях $XY_i$, $1 \le i \le p$, гиперкуба плоские сети, узлы которых имеют номера $x_{(j + iN)}$, $0 \le j \le p - 1$, и размещены на плоскостях одинаково по следам $j$. В результате мы получим *разреженный* гиперкуб с $R = Np$ узлами (рис. 3). На рис. 3 узлы на вертикальных плоскостях обозначаются как $j(x_{(j + iN)})$, а штриховые линии только выделяют вертикальные плоскости (это не ребра между узлами).

*Таблица 4*

**Размещение ПС(7, 3, 1) по плоскостям *XY* и *XZ***

| $T$ | $g$ | $XY_1$ | $XY_2$ | $XY_3$ | $g$ | $XZ_1$ | $XZ_2$ | $XZ_3$ |
|---|---|---|---|---|---|---|---|---|
| 0<br>1<br>2 | 1 | 0<br>1<br>2 | 7<br>8<br>9 | 14<br>15<br>16 | 1 | 0<br>1<br>2 | 7<br>8<br>9 | 14<br>15<br>16 |
| 3<br>4 | 2 | 3<br>4 | 10<br>11 | 17<br>18 | 2 | 17<br>18 | 3<br>4 | 10<br>11 |
| 5<br>6 | 3 | 5<br>6 | 12<br>13 | 19<br>20 | 3 | 12<br>13 | 19<br>20 | 5<br>6 |

Теперь перераспределим узлы между горизонтальными плоскостями $XZ_j$, $1 \le j \le p$, не меняя их принадлежность к вертикальным плоскостям $XY_i$ (рис. 4), так чтобы в во всех них (как $XY_i$, так и $XZ_j$) были плоские сети из $N$ узлов, т. е. с шаблоном сети ПС($N$, $p$, σ). На рис. 4 штриховые линии оконтуривают горизонтальные плоскости (это не ребра между узлами).

Процедура перераспределения узлов при σ = 1 состоит в следующем. Все узлы плоских сетей на вертикальных плоскостях $XY$ разбиваются на $p$ групп с номерами $g$. Первая группа состоит из $x$ узлов, а остальные — из $y$ узлов, где $x + y(p - 1) = N$ (левая часть табл. 4 при $p = 3$). Такое разбиение $d$ по группам будем задавать парой ($x$, $y$), которая при σ = 1 является парой ($p$, $p - 1$). Номера узлов группы $g$ в плоскостях $XY$ сдвигаются циклически на $g - 1$ ячеек вправо, образуя при этом таблицы плоских сетей на плоскостях $XZ$ (правая часть табл. 4).

При произвольном $p$ плоские сети в плоскостях $XZ$ строятся аналогично (табл. 5).

Построение плоских сетей для плоскостей $XZ$ при σ > 1 осуществляется аналогично. В табл. 6 приведен пример для $p = 4$ и σ = 2. Здесь первая группа состоит из одного узла, а остальные — из двух узлов, т. е. разбиение $d$ по группам задается парой (1, 2).

Разбиение $d$ по группам для остальных сетей ПС($N$, $p$, 2) задается табл. 7, а для сетей ПС($N$, $p$, 3) — табл. 8.

В заключение данного параграфа рассмотрим табл. 9, задающую число узлов в построенных трехмерных разреженных гиперкубах. В ней подчеркнутые значения относятся к гиперкубам, построенным на 1-расширенных плоских сетях ПС($N^*$, $p$, $\sigma|\sigma + 1$), в которых $N^* < N$. Остальные значения задаются формулой $R(p, \sigma) = N(p, \sigma)p$, т. е. $R(p, 1) = p^3 - p^2 + p$, $R(p, 2) = p^3/2 - p^2/2 + p$ и $R(p, 3) = p^3/3 - p^2/3 + p$.

*Таблица 5*

**Размещение ПС($N$, $p$, 1) по плоскостям $XZ$**

| $T$ | $g$ | $XZ_1$ | $XZ_2$ | $XZ_3$ | $XZ_4$ | ... | $XZ_p$ |
|---|---|---|---|---|---|---|---|
| 0<br>...<br>$p-1$ | 1 | 0<br>...<br>$p-1$ | $N$<br>...<br>$N+p-1$ | $2N$<br>...<br>$2N+p-1$ | $3N$<br>...<br>$3N+p-1$ | ...<br>...<br>... | $(p-1)N$<br>...<br>$(p-1)N+p-1$ |
| $p$<br>...<br>$2p-2$ | 2 | $(p-1)N+p$<br>...<br>$(p-1)N+2p-2$ | $p$<br>...<br>$2p-2$ | $N+p$<br>...<br>$N+2p-2$ | $2N+p$<br>...<br>$2N+2p-2$ | ...<br>...<br>... | $(p-2)N+p$<br>...<br>$(p-2)N+2p-2$ |
| $2p-1$<br>...<br>$3p-3$ | 3 | $(p-2)N+2p-1$<br>...<br>$(p-2)N+3p-3$ | $(p-1)N+2p-1$<br>...<br>$(p-1)N+3p-3$ | $2p-1$<br>...<br>$3p-3$ | $N+2p-1$<br>...<br>$N+3p-3$ | ...<br>...<br>... | $(p-3)N+2-1p$<br>...<br>$(p-3)N+3p-3$ |
| $3p-2$<br>...<br>$4p-4$ | 4 | $(p-3)N+3p-2$<br>...<br>$(p-3)N+4p-4$ | $(p-2)N+3p-2$<br>...<br>$(p-2)N+4p-4$ | $(p-1)N+3p-2$<br>...<br>$(p-1)N+4p-4$ | $3p-2$<br>...<br>$4p-4$ | ...<br>...<br>... | $(p-4)N+3p-2$<br>...<br>$(p-4)N+4p-4$ |
| ... | ... | ... | ... | ... | ... | ... | ... |
| $(p-1)^2+1$<br>...<br>$p^2-p$ | $p$ | $p^2+N-1$<br>...<br>$p^2-p+N$ | $p^2+2N-1$<br>...<br>$p^2-p+2N1$ | $p^2+3N-1$<br>...<br>$p^2-p+3N$ | $p^2+4N-1$<br>...<br>$p^2-p+4N$ | ...<br>...<br>... | $p^2-1$<br>...<br>$p^2-p$ |

*Таблица 6*

**Размещение ПС(7, 4, 2) по плоскостям $XY$ и $XZ$**

| $T$ | $g$ | $XY_1$ | $XY_2$ | $XY_3$ | $XY_4$ | $g$ | $XZ_1$ | $XZ_2$ | $XZ_3$ | $XZ_4$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 7 | 14 | 21 | 1 | 0 | 7 | 14 | 21 |
| 1<br>2 | 2 | 1<br>2 | 8<br>9 | 15<br>16 | 22<br>23 | 2 | 22<br>23 | 1<br>2 | 8<br>9 | 15<br>16 |
| 3<br>4 | 3 | 3<br>4 | 10<br>11 | 17<br>18 | 24<br>25 | 3 | 17<br>18 | 24<br>25 | 3<br>4 | 10<br>11 |
| 5<br>6 | 4 | 5<br>6 | 12<br>13 | 19<br>20 | 26<br>27 | 4 | 12<br>13 | 19<br>20 | 26<br>27 | 5<br>6 |

*Таблица 7*

**Размещение по группам для плоских сетей, изоморфных ПС($N$, $p$, 2)**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $p$ | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| $N$ | 7 | 11 | $\underline{15}$ | $\underline{21}$ | $\underline{27}$ | 37 | $\underline{42}$ | $\underline{51}$ | $\underline{63}$ |
| $d$ | (1, 2) | (3, 2) | (3, 2) | (3, 3) | (6, 3) | (5, 4) | (6, 4) | (1, 5) | (8, 5) |

*Таблица 8*

**Размещение по группам для плоских сетей, изоморфных ПС($N$, $p$, 3)**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $p$ | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| $N$ | 5 | $\underline{7}$ | 11 | 15 | $\underline{19}$ | $\underline{23}$ | $\underline{29}$ | $\underline{36}$ | $\underline{43}$ |
| $d$ | (2, 1) | (3, 1) | (1, 2) | (3, 2) | (5, 2) | (7, 2) | (2, 3) | (6, 3) | (10, 3) |

*Таблица 9*

**Число узлов $R(p, \sigma)$ в трехмерных разреженных гиперкубах**

| σ | *p* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1 | 21 | 44 | 105 | 186 | 273 | 456 | 657 | 910 | 1045 | 1596 |
| 2 | 12 | 28 | 55 | 90 | 147 | 216 | 333 | 420 | 561 | 756 |
| 3 | 9 | 20 | 35 | 66 | 105 | 152 | 207 | 290 | 396 | 516 |

Пути между любыми абонентами в плоскостях *XY* и *XZ* состоят из двух дуг, проходящих через промежуточный коммутатор (рис. 5) — выходная дуга *O* от абонента-источника к коммутатору и входная дуга *I* от коммутатора к абоненту-приемнику. Выходная дуга принадлежит обеим плоскостям и может обозначаться как $O_{XY}$ или $O_{XZ}$, в зависимости от контекста (способа маршрутизации). В дальнейшем употребляется обозначение $O_{XY}$.

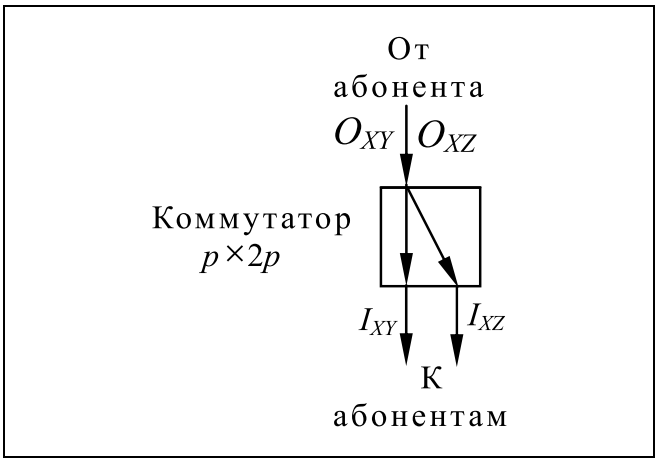


**Рис. 5. Структура коммутатора узла в плоскостях *XY* и *XZ***

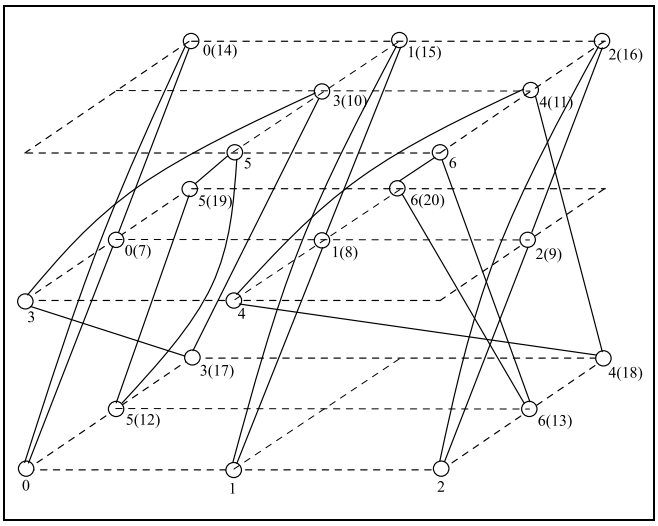


**Рис. 6. Секущие ребра $M_{YZ}$ в троичном гиперкубе**

## 3. МАРШРУТИЗАЦИЯ В ТРЕХМЕРНОМ РАЗРЕЖЕННОМ ГИПЕРКУБЕ

Построенный разреженный гиперкуб не является полносвязным, так как в нем нет путей между частью абонентов, которые находятся в разных вертикальных и горизонтальных плоскостях, например, между абонентами *0* и *11* на рис. 6.

Для преодоления этого недостатка можно дополнительно проложить ребра между узами плоскостей *XZ*, которые имеют одинаковый след (см. рис. 6). Такие ребра мы назовем секущими $M_{YZ}$, так как они пересекают плоскости *XY* и *XZ* и образуют полный граф между узлами с одинаковыми следами. При этом дуплексные каналы, соответствующие секущим ребрам, прокладываются только между коммутаторами узлов.

Теперь любой кратчайший путь, который не покрывается плоскостями *XY* и *XZ*, можно представить как трехдуговой, состоящий из дуг ребер $O_{XY}$, $M_{YZ}$ и $I_{XZ}$. Пусть узел абонента-источника находится в плоскости $XY_i$, а узел абонента-приемника — в плоскости $XZ_j$. Первое ребро $O_{XY}$ находится в плоскости $XY_i$ и проходит от абонента-источника в коммутатор первого промежуточного узла. Второе ребро $M_{YZ}$ проходит от плоскости $XY_i$ к плоскости $XZ_j$ — от коммутатора первого промежуточного узла к коммутатору второго промежуточного узла. Третье ребро $I_{XZ}$ находится в плоскости $XZ_j$ и проходит от коммутатора второго промежуточного узла к абоненту-приемнику.

Такая маршрутизация основывается на том факте, что в качестве промежуточных узлов выбираются узлы с одинаковыми следами в плоскостях $XY_i$ и $XZ_j$. Сначала находятся следы номеров узлов абонента-источника и абонента-приемника. Затем находится след промежуточного узла как при самомаршрутизации в ПС($N$, $p$, σ) (см. § 1) между

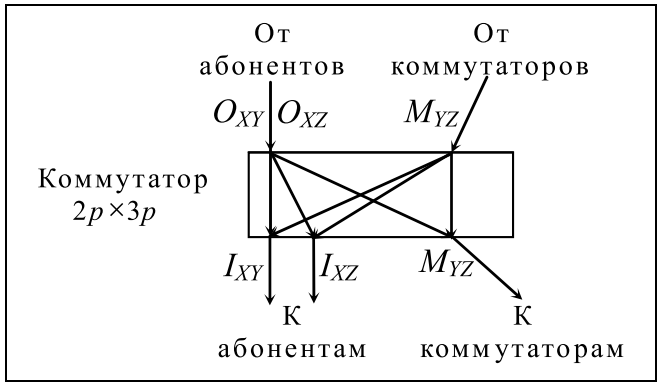


**Рис. 7. Структура коммутатора узла в разреженном гиперкубе**

источником и приемником с найденными следами. Затем по следу промежуточного узла восстанавливаются номера первого и второго промежуточных узлов в плоскостях $XY_i$ и $XZ_j$.

Рассмотрим пример маршрутизации в разреженном гиперкубе, построенном на базе ПС(7, 3, 1) (см. табл. 1 и рис. 2). Пусть узлы абонента-источника и абонента-приемника имеют номера 0 и 11 со следами 0 и 4. Они лежат (см. рис. 6) в разных плоскостях — $XY_1$ и $XZ_3$ соответственно. В сети ПС(7, 3, 1) путь от источника в узле 0 к приемнику в узле 4 осуществляется через коммутатор в узле 0. Поэтому первый и второй промежуточные узлы имеют след 0 — это узлы с номерами 0 и 14 в плоскостях $XY_1$ и $XZ_3$. Поэтому трехдуговой путь содержит внутриузловую дугу от источника к коммутатору в узле 0, дугу секущего ребра между коммутаторами в узлах 0 и 14 (рис. 7) и дугу от коммутатора в узле 14 к приемнику в узле 11.

Важно подчеркнуть, что при $\sigma > 1$ рассмотренный способ маршрутизации по следам сохраняет наличие $\sigma$ разных путей между любыми источниками и приемниками, так как они проходят через промежуточные узлы с разными следами.

Как уже показывалось в § 1, кратчайшие пути между любыми абонентом-источником и абонентом-приемником, узлы которых расположены на одной вертикальной или горизонтальной плоскости, являются двухдуговыми — от абонента-источника к промежуточному коммутатору и от него к абоненту-приемнику. Поэтому кратчайший путь между ними не содержит никаких секущих ребер. Как следствие, двухдуговые и трехдуговые пути не могут иметь конфликтов между собой — конфликтовать могут только трехдуговые пути на секущих ребрах.

В заключение данного параграфа отметим, что для прокладки кратчайших двухдуговых и трехдуговых путей коммутатор каждого узла должен иметь схему, представленную на рис. 7.

## 4. НЕБЛОКИРУЕМОСТЬ В РАЗРЕЖЕННОМ ТРЕХМЕРНОМ ГИПЕРКУБЕ

Аналогичными маршрутными свойствами обладает и 3-мерный $p$-ичный гиперкуб, построенный на базе квазиполного орграфа [3]. Главное отличие 3-мерного разреженного гиперкуба от обычного 3-мерного гиперкуба, построенного в данной работе на базе квазиполного графа $BD(N, p, \sigma)$, — это возможность иметь $\sigma$ разных кратчайших путей между любыми двумя абонентами. Однако по своим маршрутным свойствам эти гиперкубы очень близки.

В работе [3] стандартный 3-мерный $p$-ичный гиперкуб удалось сделать неблокируемым благодаря увеличению числа межузловых ребер и применению динамической локальной самомаршрутизации пакетов. Неблокируемость в стандартном 3-мерном $p$-ичном гиперкубе достигалась в работе [3] путем утроения числа вертикальных ребер измерения $Y$ и удвоения числа горизонтальных ребер измерения $Z$. Будем применять аналогичный метод и в настоящей работе с заменой вертикальных ребер на секущие ребра.

Рассмотрим кратчайший трехдуговой путь от источника в узле $s$ к приемнику в узле $d$ через промежуточные узлы $t$ и $i$ (рис. 8). Первая дуга от источника $s$ в первый промежуточный узел $t$ является дугой ребра $O_{XY}$. Через узел $t$ могут проходить пути от разных источников, которые в качестве вторых дуг используют дуги ребер $M_{YZ}$. Если каждую дугу использует только один путь через узел $i$, то он является бесконфликтным и остается таковым на последней дуге ребра $I_{XZ}$.

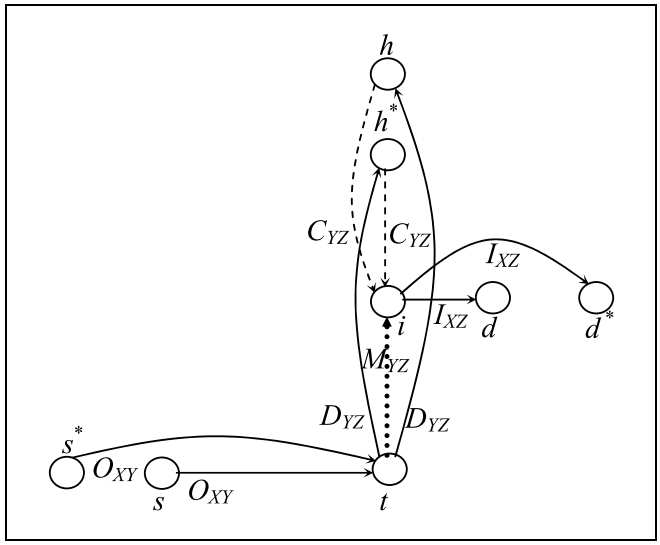


**Рис. 8. Конфликт первого типа и его разрешение**

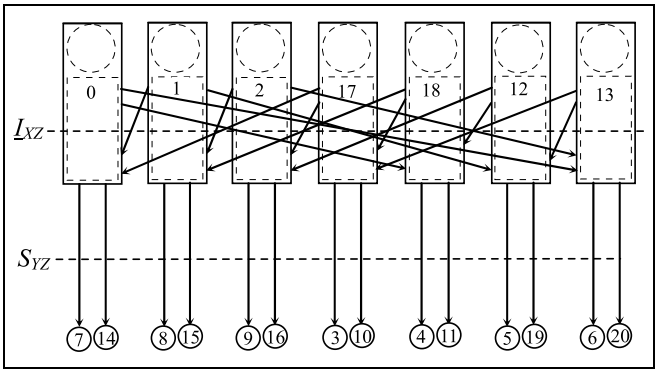


**Рис. 9. Дуги наборов $\underline{I}_{XZ}$ и $S_{YZ}$**

Будем задавать дуги ребер $O_{XY}$ и $M_{YZ}$ номерами портов $m_1$ и $m_2$ ($0 \le m_1, m_2 \le p - 1$), из которых они выходят. Для дуг $M_{YZ}$ эти порты нумеруются в порядке возрастания длин дуг, задаваемых следующим образом. Пусть дуга $M_{YZ}$ проходит из плоскости $XZ_i$ в плоскость $XZ_j$, тогда ее длина $L(i, j) = j - i$, если $(j - i) \ge 0$, и $L(i, j) = p + (j - i)$, если $(j - i) < 0$.

Ситуация, при которой на одну дугу $M_{YZ}$ претендуют несколько путей, считается конфликтом первого типа. При его возникновении один из этих путей имеет приоритет и проходит по дуге $M_{YZ}$, а остальные остаются конфликтными в узле $t$. Такой конфликт имеет место, если на прохождение через узел $t$ претендуют несколько путей (рис. 9) из разных узлов-источников $s$ и $s^*$. Здесь и далее для краткости мы обозначаем только два таких узла, которых на самом деле может быть любое число от 2 до $p - 1$. На рис. 8 конфликтная дуга обозначена полужирным пунктиром, так как она при последующем разрешении конфликта не используется.

Для разрешения конфликта необходимо использовать в каждом узле дуги еще $p - 1$ ребер $D_{YZ}$ и $C_{YZ}$, которые дублируют ребра $M_{YZ}$ и обозначены на рис. 8. Использование этих дуг позволяет избежать повторных конфликтов между дугами секущих ребер. Порты выходных дуг $M_{YZ}$, $D_{YZ}$ и $C_{YZ}$ в описании дальнейшей маршрутизации нумеруются одинаково.

Конфликт разрешается таким образом: по заведомо разным дугам $D_{YZ}$ из узла $t$ прокладываются пути в разные вспомогательные узлы $h$ и $h^*$ (без заливки на рис. 6) и осуществляется возврат из них в узел $i$ по встречным дугам ребер $C_{YZ}$. Если на дугах $C_{YZ}$ нет конфликтов, то проложенные пути бесконфликтные. Они останутся бесконфликтными и на последних дугах $I_{XZ}$ от коммутатора узла $i$ к приемникам в узлах $d$ и $d^*$.

Заведомо разными являются те дуги $D_{YZ}$, которые выходят из тех же по номерам портов, что и дуги $O_{XY}$ из узлов $s$ и $s^*$ в узел $t$. Пусть конфликтная дуга $M_{YZ}$ из узла $t$ в узел $i$ выходит из порта с номером $u$, а дуга $D_{YZ}$ — из порта с номером $v$. Тогда в качестве возвратной дуги из узла $h$ в узел $i$ выбирается дуга $C_{YZ}$ с номером $w = |v - u|$. Такая дуга всегда найдется, так как ребра $C_{YZ}$ образуют полный граф.

Однако и на дугах $C_{YZ}$ могут возникать конфликты второго типа. Каждый из них происходит в том случае, если на проход через узел $i$ на втором этапе претендуют несколько путей из разных узлов $t$ и $t^*$ после первого этапа (рис. 10) и после возникновения конфликта первого типа. Для его разрешения в каждом узле потребуется иметь еще два набора дуг. Первый набор $\underline{I}_{XZ}$ содержит дуги между коммутаторами узлов, плоских сетей на плоскостях $XZ_j$, $1 \le j \le p - 1$. Эти дуги являются копиями дуг $I_{XZ}$, но проложены между коммутаторами узлов (см. рис. 9). Второй набор $S_{YZ}$ содержит секущие дуги из коммутаторов узлов со следом $t$ к приемникам в других узлах с тем же следом (рис. 10). Дуги $S_{YZ}$ являются копиями дуг $M_{YZ}$, $D_{YZ}$ и $C_{YZ}$, но проложены от коммутаторов к абонентам.

Конфликт второго типа разрешается следующим образом (см. рис. 10). Из промежуточного узла $h$ пути прокладываются по тем дугам $\underline{I}_{XZ}$, которые равны дугам $I_{XZ}$ к приемникам. Все эти дуги $I_{XZ}$ различны, так как они завершают прямые пути из узла $i$. Поэтому различны и дуги $\underline{I}_{XZ}$, а пути про-

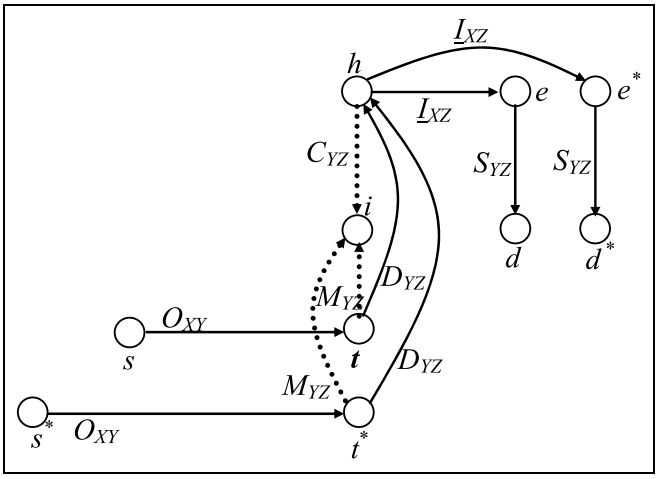


**Рис. 10. Конфликт второго типа и его разрешение**

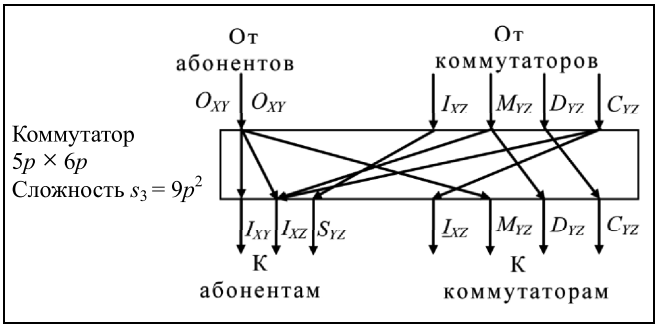


**Рис. 11. Структура коммутатора узла в неблокируемом разреженном гиперкубе**

кладываются в разные вспомогательные узлы $e$ и $e^*$, из которых они завершаются бесконфликтно по разным дугам $S_{YZ}$.

В рассмотренном разреженном гиперкубе внутриузловой коммутатор должен иметь структуру, представленную на рис. 11.

## 5. ПРОТОКОЛ САМОМАРШРУТИЗАЦИИ В РАЗРЕЖЕННОМ ГИПЕРКУБЕ

Самомаршрутизация в 3-мерном разреженном гиперкубе осуществляется как червячная маршрутизация путем прокладки бесконфликтного прямого пути между любыми источником и приемником через промежуточные коммутаторы. Она осуществляется путем посылки пилотного пакета, содержащего маршрутную информацию, с фиксацией прямого пути в промежуточных коммутаторах. Пакет данных передается по прямому пути без задержек на его буферизацию в промежуточных узлах.

Сначала представим процедуру самомаршрутизации в терминах дуг плоскостей $XY$, $XZ$ и секущих плоскость $YZ$ ребер, считая, что исходный маршрут имеет вид $O_{XY}M_{YZ}I_{XZ}$.

*Процедура динамической самомаршрутизации*

1. Если дуга $O_{XY}$ отсутствует, то путь прокладывается по локальной дуге в узле $s$ от источника к коммутатору. Переход к п. 3.

Если дуга $O_{XY}$ присутствует, то путь прокладывается от источника в узле $s$ по дуге $O_{XY}$ к коммутатору узла $t$. Переход к п. 2.

2. Если дуги $M_{YZ}$ и $I_{XZ}$ отсутствуют, то путь прокладывается по локальной дуге в узле $t$ от коммутатора к источнику. Конец алгоритма.

3. Если дуга $M_{YZ}$ отсутствует, а дуга $I_{XZ}$ присутствует, то путь прокладывается по дуге $I_{XZ}$ от коммутатора узла $t$ к приемнику в узле $d$. Выход из процедуры.

Если дуга $M_{YZ}$ присутствует, а дуга $I_{XZ}$ отсутствует (узлы $t$ и $d$ совпадают), то путь прокладывается по дуге $M_{YZ}$ от коммутатора узла $t$ к приемнику по локальной дуге. Выход из процедуры.

Если дуги $M_{YZ}$ и $I_{XZ}$ присутствуют, то проверяется наличие конфликта на дуге $M_{YZ}$. Если конфликта нет, то путь прокладывается по дуге $M_{YZ}$ от коммутатора узла $t$ к коммутатору узла $i$. Переход к п. 4.

Если на дуге $M_{YZ}$ конфликт имеет место, то путь прокладывается по уникальной для узла дуге $D_{YZ}$ (разной для разных путей) от коммутатора узла $t$ к коммутатору узла $h$. Переход к п. 5.

4. Путь прокладывается по дуге $I_{XZ}$ от коммутатора узла $i$ к абоненту в узле-приемнике $d$. Выход из процедуры.

5. Проверяется наличие конфликта на возвратной дуге $C_{YZ}$ из узла $h$.

Если конфликта нет, то путь прокладывается по дуге $C_{YZ}$ от коммутатора узла $h$ к коммутатору узла $i$. Переход к п. 4.

Если же конфликт имеет место, то переход к п. 6.

6. Путь прокладывается по дуге $\underline{I}_{XZ}$ от коммутатора узла $h$ к коммутатору узла $e$. Переход к п. 7.

7. Путь прокладывается по дуге $S_{YZ}$ от коммутатора узла $e$ к приемнику в узле $d$. Выход из процедуры.

Теперь представим алгоритм локальной самомаршрутизации в терминах путевой информации, которая содержится в пилотном пакете для червячной маршрутизации. Этот алгоритм конкретизирует процедуру самомаршрутизации, так как учитывает значения номеров портов коммутаторов, передаваемых в заголовках пакетов.

Путевую информацию будем представлять номерами портов выходных дуг в каждом узле, которые должны составлять прямой путь. Они задаются числами от 0 до $p - 1$. Номер 0 задает либо внутриузловую дугу в узле-источнике и в узле-приемнике, либо отсутствие дуги между коммутаторами разных узлов. Таким образом, выходные дуги $O_{XY}$ задаются как числа $m_1$, $0 \le m_1 \le p - 1$, секущие дуги $M_{YZ}$ задаются как числа $m_2$, $0 \le m_2 \le p - 1$, и дуги $I_{XZ}$ задаются как числа $m_3$, $0 \le m_3 \le p - 1$.

Любой кратчайший путь задается тремя числами $m_1$, $m_2$ и $m_3$. Для прокладки прямого пути пилотный пакет должен содержать пять чисел $m_1$, $m_2$, $m_3$ и $m_2^*$, $m_3^*$, где $m_2^* = m_1$, а $m_3^* = |m_2 - m_2^*|$. Эти числа вычисляются каждым источником заранее.

Прямой путь прокладывается по следующему алгоритму.

*Алгоритм динамической самомаршрутизации*

*Шаг 1.* Если $m_1 = 0$, то путь прокладывается по локальной дуге в узле абонента-источника к коммутатору. Переход к шагу 3.

Если $m_1 > 0$, то путь прокладывается от абонента-источника по дуге $O_{XY}(m_1)$. Переход к шагу 2.

*Шаг 2.* Если $m_2 = 0$ и $m_3 = 0$, то путь прокладывается по локальной дуге в узле абонента-приемника от коммутатора к абоненту. Выход из алгоритма.

*Шаг 3.* Если $m_2 = 0$ и $m_3 \neq 0$, то путь прокладывается по дуге $I_{XZ}(m_3)$ к абоненту-приемнику. Выход из алгоритма.

Если $m_2 \neq 0$ и $m_3 = 0$, то путь прокладывается по дуге $M_{YZ}(m_2)$ к коммутатору узла абонента-приемника и по локальной дуге к абоненту-приемнику. Выход из алгоритма.

Если $m_2 > 0$, то проверяется наличие конфликта на дуге $M_{YZ}(m_2)$. Если конфликта нет, то путь прокладывается по этой дуге $M_{YZ}(m_2)$. Переход к шагу 4.

Если $m_2 > 0$ и имеет место конфликт, то путь прокладывается по дуге $D_{YZ}(m_2^*)$. Переход к шагу 5.

*Шаг 4.* Путь прокладывается по дуге $I_{XZ}(m_3)$ к абоненту-приемнику. Выход из алгоритма.

*Шаг 5.* Проверяется конфликтность пути по дуге $C_{YZ}(m_3^*)$.

Если конфликта нет, то путь прокладывается по дуге $C_{YZ}(m_3^*)$. Переход к шагу 4.

Если же конфликт имеет место, то переход к шагу 6.

*Шаг 6.* Путь прокладывается по дуге $\underline{I}_{XZ}(m_3)$. Переход к шагу 7.

*Шаг 7.* Путь прокладывается по дуге $S_{YZ}(m_3^*)$. Выход из алгоритма.

Отметим, что прокладка прямых путей по приведенному алгоритму может содержать до четырех этапов. Однако передача всех пакетов данных по прямым путям осуществляется за один скачок. Последнее свойство позволяет утверждать, что неблокируемый гиперкуб практически (по задержкам передачи) имеет единичный диаметр.

## ЗАКЛЮЧЕНИЕ

Предложена новая структура трехмерного неблокируемого отказоустойчивого гиперкуба на базе плоских сетей с топологией квазиполного графа. Такой гиперкуб назван *разреженным* $p$-ичным гиперкубом, так как он содержит меньше узлов, чем неблокируемый 3-мерный $p$-ичный гиперкуб [3], но позволяет разменивать число узлов на число разных каналов между ними, что обеспечивает его канальную отказоустойчивость. Передача данных в разреженном гиперкубе осуществляется по прямым каналам между абонентами, что обеспечивает его наибольшее быстродействие. Внутриузловой коммутатор в разреженном гиперкубе имеет в 9/8 раз большее число портов, чем в неблокируемом гиперкубе или мультикольце [3, 9].

Построенный неблокируемый разреженный гиперкуб может иметь несколько сотен абонентов (процессорных ядер) при двух-трех разных прямых каналах между ними. Наличие нескольких каналов не только обеспечивает отказоустойчивость данного гиперкуба, но и открывает возможность дальнейшего повышения его быстродействия путем их параллельного применения.

Область применимости данного разреженного гиперкуба — системная сеть в многоядерном однокристальном процессоре-ускорителе.

## ЛИТЕРАТУРА


1. *Каравай М.Ф.*, *Подлазов В.С.* Метод инвариантного расширения системных сетей многопроцессорных вычислительных систем. Идеальная системная сеть // Автоматика и телемеханика. — 2010. — № 12. — С. 166—176. [*Karavay*, *M.F.*, *Podlazov*, *V.S.* An Invariant Extension Method for System Area Networks of Multicore Computational Systems. An Ideal System Network // Automation and Remote Control. — 2010. Vol. 71. — No. 12. — P. 2644—2654.]
2. *Каравай М.Ф.*, *Подлазов В.С.* Распределенный полный коммутатор как «идеальная» системная сеть для многопроцессорных вычислительных систем // Управление большими системами. — 2011. — Вып. 34. — С. 92—116. [*Karavai*, *M.F.*, *Podlazov*, *V.S.* Raspredelennyi polnyi kommutator kak «ideal'naya» sistemnaya set' dlya mnogoprotsessornykh vychislitel'nykh sistem // Upravlenie bol'shimi sistemami. — 2011. — No. 34. — S. 92—116. (In Russian)]
3. *Подлазов В.С.* Бесконфликтная самомаршрутизация для трехмерного обобщенного гиперкуба // Проблемы управления. — 2018. — No. 3. — С. 26—32. [*Podlazov*, *V.S.* Beskonfliktnaya samomarshrutizatsiya dlya trekhmernogo obobshchennogo giperkuba // Control Sciences. — 2018. — No. 3. — P. 26—32. (In Russian)]
4. *Bhuyan*, *L.N.*, *and Agrawal*, *D.P.* Generalized Hypercube and Hyperbus Structures for a Computer Network // IEEE Trans. on Computers. — 1984. — Vol. C-33, no 4. — P. 323—333.
5. *Alverson*, *R.*, *Froese*, *E.*, *Kaplan*, *L.*, *and Roweth*, *D.* Cray® XC™ Series Network. — URL: https://www.cray.com/sites/default/files/resources/CrayXCNetwork.pdf
6. *Каравай М.Ф.*, *Пархоменко П.П.*, *Подлазов В.С.* Комбинаторные методы построения двудольных однородных минимальных квазиполных графов (симметричных блок-схем) // Автоматика и телемеханика. — 2009. — № 2. — С. 153—170. [*Karavay*, *M.F.*, *Parkhomenko*, *P.P.*, *Podlazov*, *V.S.* Combinatorial methods for constructing bipartite uniform minimal quasicomplete graphs (symmetrical block designs // Automation and Remote Control. — 2009. — Vol. 70, no. 2. — P. 312—327.]
7. *Холл М.* Комбинаторика. — М.: Мир, 1970. — Гл. 10—12. [*Hall*, *M.* Combinatorial Theory. — Waltham: Blaisdell Publishing Company, 1967. — Ch. 10—12.]

8. *Каравай М.Ф.*, *Подлазов В.С.* Расширенные блок-схемы для идеальных системных сетей // Проблемы управления. — 2012. — № 4. — С. 45—51. [*Karavay*, *M.F.*, *Podlazov*, *V.S.* Rasshirennye blok-skhemy dlya ideal'nykh sistemnykh setei // Control Sciences. — 2012. — No. 4. — P. 45—51. (In Russian)]
9. *Подлазов В.С.* Бесконфликтная самомаршрутизация для трехмерного полного мультикольца // Проблемы управления. — 2018. — № 4. — С. 54—60. [*Podlazov*, *V.S.* Conflict-Free Self-Routing for a Three-Dimensional Complete Multiring // Automation and Remote Control. — 2019. — Vol. 80, no. 5. — P. 936—945.]

*Статья представлена к публикации членом редколлегии В.М. Вишневским.*

*Поступила в редакцию 9.09.20, после доработки 2.12.2019.*
*Принята к публикации 2.12.2019.*

**Подлазов Виктор Сергеевич** — д-р техн. наук,
✉ podlazov@ipu.ru,

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

# FAULT-TOLERANT NON-BLOCKING THREE-DIMENSIONAL SPARSE HYPERCUBE

V.S. Podlazov

V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia
✉ podlazov@ipu.ru

**Abstract.** A system network is proposed in the form of a non-blocking fault-tolerant three-dimensional generalized *p*-ary hypercube with a single processor in each node of the hypercube. In any non-blocking network, data between processors is transmitted over direct channels with the lowest latencies and without intermediate buffering. Networks with the topology of a generalized hypercube have the smallest lengths of direct channels and the smallest transmission delays. The structure of this hypercube based on networks with the topology of a quasi-complete graph is developed, which allows exchanging the number of processors for the number of different direct channels between any processors and setting the channel and node fault tolerance of the network. The parameters of quasi-complete graphs that exist for any *p*-identity of a hypercube are given. As a result, a network structure is proposed in the form of a sparse *p*-ary hypercube with a number of nodes slightly smaller than in a regular *p*-ary hypercube. The sparse hypercube is designed as a fault-tolerant system network for a single-chip processor accelerator with several hundred cores. A procedure and an algorithm for the laying of conflict-free direct channels through dynamic local packet self-routing has been developed, in which the nodes do not interact with each other and use only extended routing information from the packets.

**Keywords:** system network, quasi-complete graph, generalized hypercube, channel fault tolerance, switching properties, non-blocking network, conflict-free direct channels, local dynamic self-routing.

## Новые издания ИПУ РАН

**Кульба В.В., Сиротюк В.О.** Формализованная методология повышения эффективности и качества патентных информационных фондов и опыт ее использования при формировании и развитии евразийского патентно-информационного пространства. — М.: ИПУ РАН, 2019. — 236 с. — ISBN 978-5-91450-238-3.

**Сомов С.К.** Сохранность информации в распределенных системах обработки данных. — М.: ИПУ РАН. — 2019. — 254 с. — ISBN 978-5-91450-235-2.

**Лазарев А.А.** Теория расписаний. Методы и алгоритмы. — М.: ИПУ РАН, 2019. — 408 с. — ISBN 978-5-91450-236-6.

**Проблемы** управления безопасностью сложных систем: материалы XXVII Междунар. конф., 18 дек. 2019 г., Москва / под общ. ред. А.О. Калашникова, В.В. Кульбы. — М.: ИПУ РАН. — 2019. — 459 с. — ISBN 978-5-91450-241-3.