НЕБЛОКИРУЕМЫЕ ОТКАЗОУСТОЙЧИВЫЕ ДУАЛЬНЫЕ ФОТОННЫЕ КОММУТАТОРЫ ШИРОКОЙ МАСШТАБИРУЕМОСТИ

В.С. Подлазов

Аннотация. Продолжена работа по построению принципиально нового класса системных сетей – дуальных фотонных сетей, которые являются неблокируемыми сетями со статической самомаршрутизацией, обладают широкой масштабируемостью при максимальном достижимом на них быстродействии и сложностью, меньшей сложности полного коммутатора, а также допускают балансировку соотношений масштабируемость/быстродействие и сложность/быстродействие. Наконец, такие сети выполнены в расширенном схемном базисе, состоящем из дуальных фотонных коммутаторов и отдельных фотонных мультиплексоров и демультиплексоров. Предложен метод построения отказоустойчивой дуальной сети с указанными свойствами на основе сетей с топологией квазиполного графа и квазиполного орграфа и метода инвариантного расширения с внутренним распараллеливанием. Предложен метод расширения построенной ранее двухкаскадной дуальной сети в четырехкаскадную и восьмикаскадную дуальные сети с широким масштабированием при сохранении периода исходной сети и уменьшении ее показательной сложности.

Ключевые слова: фотонный коммутатор, дуальный коммутатор, фотонные мультиплексоры и демультиплексоры, многокаскадный коммутатор, бесконфликтная самомаршрутизация, неблокируемый коммутатор, статическая самомаршрутизация, квазиполный орграф, квазиполный граф, инвариантное расширение сетей, коммутационные свойства, прямые каналы, масштабируемость и быстродействие.

ВВЕДЕНИЕ

Данная работа развивает методику построения принципиально нового класса системных сетей [1– 6] – дуальных фотонных сетей, которые являются неблокируемыми сетями со статической самомаршрутизацией [1–3, 5] и могут иметь заданную степень канальной отказоустойчивости [6].

В работе предложена методика построения неблокируемых самомаршрутизируемых фотонных сетей широкой масштабируемости. Это так называемые дуальные сети, в основе которых лежит неблокируемый дуальный коммутатор $p \times p$ с периодом разрядов в p длительностей сигнала (тактов) [1–3]. Дуальный коммутатор совмещает шинный способ разрешения конфликтов сигналов (разведение по разным тактам в одном канале) и коммутаторный способ (разведение по разным каналам). Дуальный коммутатор является неблокируемым коммутатором на любом входном трафике, если информационные разряды передаются с периодом в *p* тактов. Дуальный коммутатор был разработан коллегами автора [1, 2], а затем применен и поименован в совместных работах [3–5]. Фактически его использование оказалось необходимым условием построения неблокируемых сетей широкой масштабируемости и приемлемой сложности.

В дуальном фотонном коммутаторе применяется способ передачи сигнальной и управляющей информации параллельно на разных частотах для каждого разряда данных. Этот способ позволяет снять проблему синхронизации сигналов разных каналов.

Фотонная специфичность рассматриваемых сетей состоит в использовании внутриразрядной виртуализации каналов посредством обратных связей через линии задержки длительностью в один такт и, кроме того, в использовании управляющих сигналов на разных частотах для выполнения маршрутизации отдельных разрядов. Разведение информационных сигналов по разным тактам со-



провождается разведением по тем же тактам и сопровождающей управляющей информации. По ней осуществляется маршрутизация разрядов посредством перемещения их между разными каналами без изменения установившихся номеров тактов.

По всей статье термин «дуальный коммутатор» или «дуальная сеть, построенная с использованием дуальных коммутаторов», подразумевает использование в них разрядов с периодом в *р* тактов. Эти разряды обеспечивают неблокируемость сети в ее первом каскаде, а в остальных каскадах остаются по «инерции», уже без применения шинного способа разрешения конфликтов.

Масштабируемость дуальных сетей осуществляется с использованием сетей с топологией квазиполных графа или орграфа [4], которые выполняются в расширенном схемном базисе, состоящем из дуальных фотонных коммутаторов и отдельных фотонных мультиплексоров и демультиплексоров. В работах [1–3, 5] широкая масштабируемость достигается посредством применения метода инвариантного расширения сетей с большим числом дополнительных демультиплексоров и мультиплексоров.

В статьях [5, 6] был разработан и впервые применен новый метод расширения дуальных сетей посредством их внутреннего распараллеливания без использования дополнительных устройств. В частности, в работе [5] была построена двухкаскадная неблокируемая сеть, которая состоит в каждом каскаде из сетей с топологией квазиполного орграфа, имеющих $N = p^2$ каналов, тогда как двухкаскадная неблокируемая сеть имеет N^2 каналов. С другой стороны, в работе [6] была построена двухкаскадная неблокируемая сеть с (σ – 1)канальной отказоустойчивостью, которая состоит в каждом каскаде из сетей с топологией квазиполного графа, имеющих $N = p(p-1)/\sigma + 1$ каналов, а двухкаскадная отказоустойчивая неблокируемая сеть имеет N^2 каналов.

В настоящей работе строятся четырех- и восьмикаскадные отказоустойчивые сети посредством развития и применения обобщенного метода внутреннего распараллеливания. При этом достигается такая же степень масштабирования сетей, как и в инвариантном методе с использованием внешних демультиплексоров и мультиплексоров, но без их применения, и строятся сети с существенно меньшей сложностью.

В § 1 дается обзор наличия свойств неблокируемости и канальной отказоустойчивости у ряда современных системных сетей. В § 2 вслед за статьей [6] вводятся понятия прямой и обратной *р*перестановки, которые существенно используются при доказательстве неблокируемости четырех- и восьмикаскадных сетей, и повторяются доказательства неблокируемости двухкаскадных сетей. В § 3 данной части строятся четырехкаскадные неблокируемые самомаршрутизируемые коммутаторы с одноканальной и двухканальной отказоустойчивостью и оцениваются их характеристики. При этом метод внутреннего распараллеливания из работ [5, 6] обобщается для четырехкаскадных коммутаторов.

В § 4 сравниваются характеристики четырехкаскадных неблокируемых самомаршрутизируемых коммутаторов на базе коммутаторов с топологией дуальных квазиполных графов и орграфов. Наконец, в § 5 строятся восьмикаскадные неблокируемые самомаршрутизируемые коммутаторы на базе коммутаторов с топологией дуальных квазиполных графов и орграфов. При этом метод из § 3 обобщается для восьмикаскадных коммутаторов. В § 6 обсуждаются свойства построенных сетей в сравнении с другими неблокируемыми сетями и недостатки построенных сетей, а также намечаются способы преодоления этих недостатков.

В заключении дается разбор сути обобщенного метода внутреннего распараллеливания, который представляет собой ядро методики построения дуальных неблокируемых сетей широкой масштабируемости и невысокой удельной сложности. Здесь выделяются три основных компонента предложенной методики – неблокируемый дуальный коммутатор, коммутатор с топологией квазиполного графа или орграфа на базе дуального коммутатора и метод внутреннего распараллеливания.

1. НЕБЛОКИРУЕМОСТЬ И ОТКАЗОУСТОЙЧИВОСТЬ В СИСТЕМНЫХ СЕТЯХ

Задача построения неблокируемых отказоустойчивых системных сетей суперкомпьютеров до настоящего времени не имеет полного решения.

Системная сеть является неблокируемой, если в ней для любой перестановки пакетов можно проложить бесконфликтные пути от источников к приемникам. Системная сеть является самомаршрутизируемой, если бесконфликтные пути можно проложить локально по узлам сети без их взаимодействия только на основе маршрутной информации в пакетах. Наконец, самомаршрутизация является статической, если любой источник может самостоятельно наметить бесконфликтные пути к своему приемнику без взаимодействия с другими источниками.

Существование неблокируемых сетей было доказано еще Клозом [7, 8]. Пока еще не построены процедуры самомаршрутизации неблокируемых сетей Клоза, которые, однако, могут быть мерилом качества других неблокируемых сетей. Неблокируемой является сеть в виде двумерного обобщенного гиперкуба с топологией квазиполного орграфа, например, в коммутаторах YARK и ROSETTA, использованных в ряде сетей разной структуры: перестраиваемой сети Клоза [9], трехмерного тора [10], иерархии полных и квазиполных орграфов [11–13]. К сожалению, квазиполный орграф имеет малое число каналов $N = p^2$, где p – это степень внутренних коммутаторов, и большую коммутационную сложность $S \ge N^2$, которая больше сложности полного орграфа и заведомо больше сложности неблокируемой сети Клоза.

В настоящее время в литературе широко представлены системные сети со структурой толстого дерева (в частности, перестраиваемые сети Клоза), со структурой обобщенного гиперкуба, со структурой многомерного тора и сети с иерархией полных и квазиполных орграфов.

Сети со структурой толстого дерева являются перестраиваемыми сетями [9, 14, 15], в которых бесконфликтная передача осуществляется только по заранее составленным расписаниям для конкретных перестановок пакетов. Для произвольных перестановок эти сети оказываются блокируемыми, в них произвольная перестановка полностью осуществляется за несколько скачков между узлами сети. Максимальное число таких скачков задает диаметр сети. В перестраиваемых сетях Клоза диаметр равен числу каскадов сети.

Сети со структурой обобщенного гиперкуба [16–19] не являются даже перестраиваемыми сетями [20, 21]. Их можно сделать перестраиваемыми посредством увеличения числа каналов в некоторых измерениях. Обобщенные кубы имеют диаметр, равный числу измерений или на единицу меньше в расширенном гиперкубе [17, 18]. Обобщенные гиперкубы с удвоенным числом каналов в каждом измерении являются перестраиваемыми сетями сразу для двух перестановок. Отметим, что попытка использования обобщенного гиперкуба в качестве неблокируемой сети для фотонного компьютера [22] представляется очень сомнительной затеей.

Сети со структурой многомерного тора для произвольных перестановок вообще не имеют возможности передавать пакеты по прямым каналам [11, 23–25]. В них произвольные перестановки осуществляются только за несколько скачков между узлами сети. Многомерные торы являются самыми простыми, но и самыми медленными сетями вследствие их больших диаметров. Так, в сетях, рассмотренных в работах [11, 23–25], диаметр измеряется десятками скачков.

Наоборот, сети со структурой иерархии полных или квазиполных орграфов [10, 12, 26] имеют са-

мый маленький диаметр в три скачка. Сетей с малыми диаметрами в последнее время появилось довольно много [27–32]. Все они имеют серьезные проблемы с выравниванием загрузки сети при отказах каналов.

Канальной отказоустойчивостью считается возможность сохранения полнодоступности сети при отказах каналов с сохранением ее исходных характеристик (неблокируемость сети, задержки передачи или диаметра сети).

В чистом виде канальной отказоустойчивостью обладают, по-видимому, только сети с топологией квазиполных графов. Эти сети изоморфны такому математическому объекту, как неполная уравновешенная симметричная блок-схема [33]. Эти сети имеют элементную базу из коммутаторов $p \times p$, демультиплексоров $1 \times p$ и мультиплексоров $p \times 1$ и являются неблокируемыми сетями со статической самомаршрутизацией. Они имеют прямые каналы между $N = p(p-1)/\sigma + 1$ абонентами сети и имеют о разных каналов между любыми двумя абонентами [4].

В других сетях восстановление полнодоступности сети при отказах каналов сопровождается так или иначе увеличением задержек передачи по сети. Так, в перестраиваемой сети Клоза при отказах каналов нагрузка на оставшиеся каналы увеличивается, что, в свою очередь, увеличивает число конфликтов и задержки передачи части пакетов.

Сеть TOFY со структурой трехмерного тора [25] использует еще три измерения для создания резервных каналов. В ней при отказе части колец сети их целостность восстанавливается с увеличением диаметра сети на единицу.

Обобщенные гиперкубы с удвоенным числом каналов в каждом измерении являются 1-отказоустойчивыми сетями с неизменным диаметром [19].

В сетях с иерархией полных или квазиполных орграфов [10, 12, 26] при отказе части каналов полнодоступность сети восстанавливается путем использования обходных путей с длительностью в пять скачков, т. е. с увеличением задержки передачи в 5/3 раза.

2. ДУАЛЬНЫЙ КВАЗИПОЛНЫЙ ГРАФ, РПЕРЕСТАНОВКИ И ДУАЛЬНЫЙ ДВУХКАСКАДНЫЙ Коммутатор

Дуальный коммутатор $K\Pi N_1$ с топологией квазиполного графа $K\Pi\Gamma(N_1, p, \sigma)$ состоит из $N_1 = p(p-1)/\sigma + 1$ дуальных коммутаторов $p \times p$, обозначаемых как ДKp, N_1 входных демультиплексоров $1 \times p$ и N_1 выходных мультиплексоров $p \times 1$ [6]. Схема их межсоединений строится комбинатор-

Pr-

Таблииа 2

ным методом [4]. В нем между любыми входом и выходом имеются σ разных путей через разные дуальные комутаторы ДК*p*. На рис. 1 приводится схема коммутатора КП4 как графа КПГ(4, 3, 2) с одноканальной отказоустойчивостью. Для примера в нем выделены по два пути, соединяющие два случайно выбранных входа и выхода – (2, 4) и (3, 3).



Рис. 1. Дуальный с периодом разрядов в три такта квазиполный коммутатор КП4 в виде графа КПГ(4, 3, 2): a - в исходном виде с дуплексными каналами, $\delta - в$ применяемом виде с симплексными каналами; пунктирами и точками изображены разные пути между выбранными входами и выходами

Таблица 1

Характеристики дуальных коммутаторов КП*N*₁ с одноканальной отказоустойчивостью

р	N_1	T_1	S_1	L_1
2	2	2	$24 = N_1^{4,58}$	$8 = N_1^{3}$
4	7	4	$280 = N_1^{2,9}$	$56 = N_1^{2,07}$
6	15	6	$1\ 260 = N_1^{2,64}$	$180 = N_1^{1,92}$
8	27	8	$3\ 888 = N_1^{2,51}$	$432 = N_1^{1,84}$

Любой дуальный коммутатор $K\Pi N_1$ имеет тот же период разрядов T_1 , что и входящий в него дуальный коммутатор ДКр. Для коммутатора $K\Pi N_1$ рассчитываются такие характеристики, как коммутационная сложность S_1 , выраженная в числе точек коммутации, и канальная сложность, выраженная в числе каналов. Они выражаются в показательном виде, через число каналов, и именуются в статье показательными сложностями¹ (см. табл. 1).

Для дуального коммутатора $K\Pi N_1$ с топологией квазиполного *орграфа* [5] характеристики задаются табл. 2, которые лучше чем в табл. 1, но без обеспечения канальной отказоустойчивости.

Введем понятие *p*-разбиения пакетов, передаваемых через заданное сечение сети на входах в мультиплексоры. Все пакеты разделяются на группы переменного состава, содержащие не более *p*

Характеристики дуальных коммутаторов КП <i>N</i> ₁	
на основе квазиполного орграфа	

р	N_1	T_1	S_1	L_1
2	4	2	$48 = N_1^{2,79}$	$16 = N_1^2$
4	16	4	$640 = N_1^{2,33}$	$128 = N_1^{1,75}$
6	36	6	$3\ 024 = N_1^{2,24}$	$432 = N_1^{1,69}$
8	64	8	9 216 = $N_1^{2,19}$	$1\ 024 = N_1^{1,67}$

пакетов в каждой. При обычной перестановке пакетов на входе и выходе коммутатора имеет место 1-разбиение. Назовем *p*-перестановкой вариант передачи, при котором на входе сети имеет место 1-разбиение, а на заданном сечении – *p*-разбиение.

Для дуального коммутатора КП N_1 с топологией дуального квазиполного графа КПГ (N_1, p, σ) указанное сечение проводится через входы мультиплексоров и называется выходным сечением. На рис. 1 выходное сечение обозначено вертикальным пунктиром. По свойству дуального коммутатора ДКp на выходном сечении при любом трафике имеет место p-разбиение. Имеет место

Лемма 1 [6]. Дуальный коммутатор КПN₁ с периодом разрядов р тактов является неблокируемым коммутатором со статической самомаршрутизацией на любой обычной перестановке и обладает при этом (σ – 1)-канальной отказоустойчивостью.

В работах [5, 6] была развита методика построения двухкаскадных неблокируемых коммутаторов с N_1^2 каналами. Здесь она излагается на примере дуального коммутатора КП2 с топологией дуального квазиполного графа КПГ(2, 2, 2), представленного на рис. 2.



Рис. 2. Дуальный неблокируемый коммутатор КП2 с одноканальной отказоустойчивостью

На его основе строится четырехканальная двухкаскадная сеть C_24 , которая содержит по два коммутатора КП2 в каждом каскаде, соединенных обменными связями (рис. 3). Эта сеть является блокируемой на мультиплексорах первого каскада, выделенных заливкой, и на них же утрачивает канальную отказоустойчивость.

Сеть C₂4 может быть превращена в неблокируемый коммутатор K₂4 с одноканальной отказоустойчивостью посредством метода внутреннего



¹ Термин автора.



Рис. 3. Дуальная двухкаскадная блокируемая сеть C₂4 с обменными связями

распараллеливания [5, 6]. В нем во втором каскаде образуются две копии второго каскада сети C_24 . В первом каскаде мультиплексоры на разрезах удаляются, а их входы подсоединяются к входам копий второго каскада: нечетные – к первой копии, а четные – ко второй копии. Эти соединения сохраняют порядок подсоединения тех каналов, которые в сети C_24 шли ко второму каскаду. Сами вырезанные мультиплексоры перемещаются для объединения выходов копий второго каскада, образуя выходные мультиплексоры коммутатора K_24 (рис. 4), который оказывается неблокируемым коммутатором со статической самомаршрутизацией и одноканальной отказоустойчивостью.



Рис. 4. Дуальный неблокируемый самомаршрутизируемый коммутатор К₂4 с одноканальной отказоустойчивостью

В общем случае ($p \ge 2$) дуальный коммутатор КП N_1 имеет топологию дуального квазиполного графа КПГ(N_1 , p, σ) с периодом разрядов в p тактов. На его основе строится двухкаскадная блокируемая сеть C_2N_2 ($N_2 = N_1^2$), в каждом каскаде ко-

торой содержится N_1 коммутаторов КП N_1 с обменными связями между каскадами. Для применения метода внутреннего распараллеливания образуется p копий второго каскада сети C_2N_2 , а мультиплексоры первого каскада используются для объединения одноименных выходов копий второго каскада.

На разрезах первого каскада имеется pN_2 входов в мультиплексоры. Они перенумеровываются сверху вниз как I ($1 \le I \le pN_2$), и входы с номерами i = I(modp)+1 подсоединяются к одноименным входам i-й копии второго каскада, сохраняя при этом порядок размещения по коммутаторам K_1N_1 , как в сети C_2N_2 .

Лемма 2. Дуальный коммутатор K₂N₂ имеет р-перестановку на заданном сечении. Он является неблокируемым коммутатором со статической маршрутизацией на любой обычной перестановке и обладает (σ – 1)-канальной отказоустойчивостью.

Коммутационная сложность коммутатора K_2N_2 задается по построению по рекуррентным формулам как $S_2 = N_1S_1 + pN_1S_1$, а канальная — как $L_2 = N_1L_1 + pN_1L_1$. Численные значения характеристик коммутаторов K_2N_2 для $\sigma = 2$ представлены в табл. 3. Отметим, что имеет место уменьшение значений показательных сложностей по сравнению с табл. 1.

Таблица 3

Характеристики дуальных коммутаторов К₂*N*₂ с одноканальной отказоустойчивостью

р	N_1	$N_2 = N_1^2$	$T_2 = p$	S_2	L_2
2	2	4	2	$N_2^{3,58}$	$N_2^{2,9}$
4	7	49	4	$N_2^{2,37}$	$N_2^{1,97}$
6	15	225	6	$N_2^{2,18}$	$N_2^{1,84}$
8	27	729	8	$N_2^{2,09}$	$N_2^{1,77}$

При построении дуального коммутатора K_2N_2 на базе коммутатора с топологией квазиполного орграфа [5] его характеристики задаются табл. 4, которые существенно лучше, чем в табл. 2, но без обеспечения канальной отказоустойчивости.

Таблица 4

Характеристики коммутаторов К₂ *N*₂ на основе квазиполного орграфа

р	N_1	$N_2 = N_1^2$	$T_2 = p$	S_2	L_2
2	4	16	2	$N_2^{2,29}$	$N_2^{1,95}$
4	16	256	4	$N_2^{1,96}$	$N_2^{1,68}$
6	36	1 296	6	$N_2^{1,89}$	$N_2^{1,63}$
8	64	4 096	8	$N_2^{1,86}$	$N_2^{1,6}$



Отметим также, что дуальный коммутатор K_2N_2 имеет два каскада выходных мультиплексоров, содержащих pN_2 и N_2 мультиплексоров соответственно. Для целей следующего раздела в коммутаторе K_2N_2 проводится новый штрихпунктирный разрез (см. рис. 4), проходящий через входы первого каскада мультиплексоров.

3. ЧЕТЫРЕХКАСКАДНЫЙ ОТКАЗОУСТОЙЧИВЫЙ НЕБЛОКИРУЕМЫЙ САМОМАРШРУТИЗИРУЕМЫЙ КОММУТАТОР С ДВУМЕРНЫМ ВНУТРЕННИМ РАСПАРАЛЛЕЛИВАНИЕМ

В работах [1–3, 5] дальнейшее увеличение числа каналов неблокируемого коммутатора осуществлялось с помощью метода инвариантного расширения сетей с внешним распараллеливанием, при котором не происходит изменения периода сигналов. Однако он оказался малоприменим для отказоустойчивых неблокируемых коммутаторов [6].

В данном разделе увеличение числа каналов без изменения периода сигналов осуществляется с помощью обобщенного метода внутреннего распараллеливания сети посредством построения четырехкаскадных коммутаторов K_4N_4 из двухкаскадных коммутаторов K_2N_2 с числом каналов $N_4 = N_2^2$ и периодом разрядов $T_4 = T_2 = p$.

Построение сети осуществляется на примере использования коммутаторов K_24 (см. рис. 4). Сначала создается двухкаскадная сеть C_416 . Каждый ее каскад состоит из четырех копий коммутатора K_24 , а каскады соединены обменными связями (рис. 5).

В реальности сеть C₄16 состоит из четырех каскадов КПГ(2, 2, 2), что задает нижний индекс в ее обозначении. Сеть C₄16 является блокируемой вследствие возможности возникновения конфликтов сигналов в двух каскадах выходных мультиплексоров M2, обозначенных заливкой. Имеется два слоя этих мультиплексоров общим числом $W_4 = 48$. Кроме того, в них нарушается и канальная отказоустойчивость. Через входы в первый каскад проводится штрихпунктирный разрез (см. рис. 5), для которого в предыдущем разделе сформулировано понятие *p*-перестановки.

Затем создается сеть <u>C</u>₄16, которая содержит первый каскад сети C₄16 и две копии второго каскада сети C₄16. В сети <u>C</u>₄16 создается одна параллельная схема первого измерения (рис. 6). Для этого сначала вырезается внешний слой со светлой заливкой общим числом $W_2^* = 16$ мультиплексоров M2. Они остаются пока без подсоединений. Затем вырезаются мультиплексоры M2 внутреннего слоя с темной заливкой и их нечетные входы разводятся по входам двух копий второго каскада сети C_416 . При этом $W_{4,1} = 16$ вырезанных мультиплексоров объединяют выходы этих двух копий.

Оставшиеся $W_{4,2} = 16$ мультиплексоров M2 с темной заливкой используются для создания второй параллельной схемы первого измерения аналогичным образом (рис. 7). Для этого их четные входы разводятся по входам еще двух копий второго каскада сети C₄16.

Забегая вперед, отметим, что на рис. 6 и 7 обозначены и новые соединения мультиплексоров первого и второго слоев. Они задают объединение схем первого измерения в схему второго измерения.

В результате построены две схемы первого измерения, каждая из которых состоит из двух коммутаторов K_24 , включенных параллельно (рис. 8 без соединений второй схемы первого измерения). Две схемы первого измерения образуют двумерную схему. Выходы двумерной схемы объединяют $W_2^* = 16$ мультиплексоров со светлой заливкой, образуя выходы коммутатора K_416 . На рис. 8 последние соединения представлены точечными линиями и полностью только в одном экземпляре изза недостатка площади рисунка (они есть на рис. 6 и 7).



Рис. 5. Блокируемая дуальная сеть C416 без канальной отказоустойчивости



Рис. 6. Построение первой схемы первого измерения. Неиспользованные в этой схеме мультиплексоры сети С416 приведены слева



Рис. 7. Построение второй схемы первого измерения посредством использования мультиплексоров, которые не использовались при построении первой схемы первого измерения

Ş



Рис. 8. Неблокируемый дуальный коммутатор К416 с одноканальной отказоустойчивостью

Ş

Построенная таким образом 16-канальная сеть состоит из 16 копий коммутатора K_24 , включенных параллельно. На их входы поступают разреженные альтернативные (не пересекающиеся ни по входам, ни по выходам) прямые *p*-перестановки, которые по лемме 2 реализуются бесконфликтно по единому статическому расписанию. Пути между источниками и приемниками в коммутаторах K_24 проходят по двум путям через разные схемы первого измерения. Поэтому в коммутаторе K_24 имеет место одноканальная отказоустойчивость, так как $p = \sigma = 2$.

В общем случае (p > 2 и $\sigma \ge 2$) сначала строится сеть C₄N₄, которая состоит из двух каскадов по N₂ коммутаторов K₂N₂ в каждом, соединенных обменными связями. Она имеет N₄ = N₂² каналов и допускает возможность конфликта сигналов на выходных мультиплексорах Mp. Поэтому она является блокируемой сетью и не обладает канальной отказоустойчивостью.

Первый каскад сети C_4N_4 имеет два слоя выходных мультиплексоров Mp общим числом $W_4 = N_2V_2 = N_4(p+1)$. Первый внутренний слой мультиплексоров Mp содержит $W_{4,1} = pN_4$ мультиплексоров Mp, которые в совокупности имеют p^2N_4 входов.

Затем создается сеть $\underline{C}_4 N_4$, которая содержит первый каскад сети $C_4 N_4$ и p^2 копий второго каскада сети $C_4 N_4$. Сеть $\underline{C}_4 N_4$ создается с соблюдением следующей структуры: копии второго каскада сети $C_4 N_4$ образуют p схем первого измерения, а все вместе они образуют схему второго измерения.

Сеть <u>C</u>₄ N_4 содержит p^2N_2 коммутаторов K₂ N_2 , имеющих в совокупности p^2N_4 входов. Далее вырезаются оба слоя мультиплексоров Мp первого каскада сети C₄ N_4 , а входы первого слоя мультиплексоров Мp подсоединяются к входам коммутаторов K₂ N_2 . Это возможно, так как число этих входов совпадает.

Копии второго каскада сети C_4N_4 в сети \underline{C}_4N_4 числом p^2 разобьем на p групп и обозначим через G ($1 \le G \le p$) номер группы и через I ($1 \le I \le p$) – номер копии. Фактически G задает номер схемы первого измерения, а I – номер копии в схеме первого измерения.

Обозначим через J $(1 \le J \le N_2)$ номер коммутатора K_2N_2 в схеме первого измерения. Обозначим через K $(1 \le K \le N_2)$ номер входа каждого такого коммутатора K_2N_2 . Таким образом, вход любого коммутатора K_2N_2 задается составным индексом G, I, J, K.

Обозначим через $i (0 \le i \le p - 1)$ номер входа по тоdp каждого мультиплексора Мp в первом их слое. В каждом коммутаторе первого каскада сети C_4N_4 содержится pN_2 этих мультиплексоров Mp. Разобьем их на N_2 групп и обозначим через g $(1 \le g \le N_2)$ номер группы, а через j $(1 \le j \le p)$ номер мультиплексора Mp в группе. Обозначим через k $(1 \le k \le N_2)$ номер коммутатора K_2N_2 в первом каскаде сети C_4N_4 . Таким образом, вход любого мультиплексора Mp в первом слое задается составным индексом i, g, j, k.

Произвольный вход мультиплексора Mp с индексом *i*, *g*, *j*, *k* переключается на вход коммутатора K_2N_2 с индексом *G*, *I*, *J*, *K*, в котором G = i + 1, I = j, J = g и K = k. В результате образуется p^2N_2 параллельно включенных коммутаторов K_2N_2 , на входы которых поступают разреженные непересекающиеся прямые *p*-перестановки.

Вырезанные мультиплексоры Мp первого слоя объединяют выходы схем первого измерения. Вырезанные мультиплексоры Мp второго слоя объединяют выходы p схем первого измерения, образуя выходы схемы второго уровня, которые также являются выходами и коммутатора K_4N_4 .

Для коммутаторов К₄N₄ справедлива

Лемма 3. Дуальный коммутатор K₄N₄ является неблокируемым коммутатором со статической маршрутизацией на любой обычной перестановке при любом р. Он обладает ($\sigma - 1$)-канальной отказоустойчивостью.

Д о к а з а т е л ь с т в о. Первое утверждение опирается на использование коммутатора K_2N_2 и справедливость для него леммы 2. Второе утверждение опирается на неблокируемость коммутатора K_2N_2 и тот факт, что *p*-перестановка на сечении состоит из разреженных 1-перестановок, разведенных по разным каналам и тактам.

Согласно назначению G = i+1, разные входы одного мультиплексора первого слоя подсоединяются к разным одномерным схемам (разным копиям второго каскада сети C_4N_4), а входы разных мультиплексоров – к входам разных коммутаторов K_2N_2 второго каскада сети C_4N_4 . Иначе говоря, имеется σ разных путей коммутаторов K_2N_2 по p разным путям в коммутаторе K_4N_4 , откуда следует его ($\sigma - 1$)-канальная отказоустойчивость, так как $p \ge \sigma$.

В результате неблокируемый самомаршрутизируемый коммутатор K_4N_4 с $N_4 = N_2^2$ каналами и (σ -1)-канальной отказоустойчивостью имеет характеристики, приведенные в табл. 5 и табл. 6. Они рассчитываются по рекуррентным формулам $S_4 = N_2S_2 + p^2N_2S_2$ и $L_4 = N_2L_2 + p^2N_2L_2$. Коммутатор K_4N_4 имеет четыре слоя выходных мультиплексоров М*p* общим числом $V_4 = N_4(p^4 - 1)/(p - 1)$.

Отметим дальнейшее снижение показательных коммутационной и канальной сложностей комму-

татора K_4N_4 (см. табл. 4) по сравнению с коммутатором K_2N_2 (см. табл. 2).

Если отказаться от требования канальной отказоустойчивости, то на основе коммутатора с топологией квазиполного орграфа можно построить неблокируемый самомаршрутизируемый коммутатор K_4N_4 с характеристиками, заданными в табл. 7. Видно, что по сравнению с отказоустойчивыми вариантами он имеет большее число каналов и меньшую сложность. Кроме того, весь набор коммутаторов имеет коммутационную и канальную сложности меньше, чем у двухкаскадного коммутатора на базе коммутатора с топологией полного графа (ср. с табл. 3) и меньше, чем у коммутатора с топологией полного графа (*switchboard*).

4. ВОСЬМИКАСКАДНЫЕ НЕБЛОКИРУЕМЫЕ КОММУТАТОРЫ С ЧЕТЫРЕХМЕРНЫМ ВНУТРЕННИМ РАСПАРАЛЛЕЛИВАНИЕМ НА ОСНОВЕ ГРАФА И ОРГРАФА

Метод расширения двухкаскадных коммутаторов K_2N_2 в четырехкаскадные коммутаторы K_4N_4 можно обобщить и на построение восьмикаскадных коммутаторов K_8N_8 из четырехкаскадных

коммутаторов К₄N₄.

Сначала строится сеть C_8N_8 , которая состоит из двух каскадов по N_4 коммутаторов K_4N_4 в каждом, соединенных обменными связями. Она имеет $N_8 = N_4^2$ каналов и допускает возможность конфликта сигналов на выходных мультиплексорах M_p первого каскада. Поэтому она является блокируемой сетью и не обладает канальной отказоустойчивостью.

Первый каскад сети C_8N_8 имеет четыре слоя выходных мультиплексоров Mp общим числом $W_8 = N_4V_4 = N_8(p^4 - 1)/(p - 1)$. Первый внутренний слой мультиплексоров Mp содержит $W_{8,1} = p^3N_8$ мультиплексоров Mp, которые в совокупности имеют p^4N_8 входов.

Затем создается сеть $\underline{C}_8 N_8$, которая содержит p^4 копий второго каскада сети $C_8 N_8$. Сеть $\underline{C}_8 N_8$ создается с соблюдением следующей структуры. Копии второго каскада сети $C_8 N_8$ общим числом p образу-

Таблица 7

p	N_1	$N_2 = N_1^2 = p^4$	$N_4 = N_2^4 = p^{\circ}$	$T_4 = p$	S_4	L_4
2	4	16	256	2	43 520 = $N_4^{1,93}$	$17\ 920 = N_4^{1,77}$
3	9	81	6 561	3	$6 \ 101 \ 730 = N_4^{1,78}$	$1\ 771\ 470 = N_4^{1,64}$
4	16	256	65 536	4	$2,18\mathrm{E} + 08 = N_4^{1,73}$	$49\ 020\ 928 = N_4^{1,60}$
5	25	625	390 625	5	$3,61\mathrm{E} + 09 = N_4^{1,71}$	$6,6\mathrm{E} + 08 = N_4^{1,58}$
6	36	1 296	1 679 616	6	$3,62\mathrm{E} + 10 = N_4^{1,70}$	$5,59\mathrm{E} + 09 = N_4^{1,57}$
7	49	2 401	5 764 801	7	$2,56\mathrm{E}+11 = N_4^{1,69}$	$3,43\mathrm{E}+10=N_4^{1,56}$
8	64	4 096	16 777 216	8	$1.4E + 12 = N_4^{1.68}$	$1.66E + 11 = N_4^{1.55}$

Характеристики дуальных коммутаторов К₄ N₄ на основе орграфа

Характеристики дуальных коммутаторов К₄ *N*₄ с одноканальной отказоустойчивостью

р	N_1	$N_4 = N_1^4$	$T_4 = p$	S_4	L_4
2	2	16	2	$2\ 720 = N_4^{2,85}$	$1\ 120 = N_4^{2,53}$
3	4	256	3	$238\ 080 = N_4^{2,23}$	$69\ 120 = N_4^{2,01}$
4	7	2 401	4	$8\ 000\ 132 = N_4^{2,04}$	$1\ 795\ 948 = N_4^{1,85}$
5	11	14 641	5	$1,35E+08 = N_4^{1,95}$	$24\ 743\ 290 = N_4^{1,77}$
6	15	65 536	6	$1,41E+09 = N_4^{1,9}$	$2,18E+08 = N_4^{1,73}$
7	21	194 481	7	$8,64E+09 = N_4^{1,88}$	$1,16E+09 = N_4^{1,71}$
8	27	531 441	8	$4,45E+10 = N_4^{1,86}$	$5,25E+09 = N_4^{1,7}$

Характеристики дуальных коммутаторов К₄ N₄ с двухканальной

отказоустойчивостью

 S_4

75 330 = $N_4^{2,56}$

 $2\ 082\ 500 = N_4^{2,26}$

22 161 230 = $N_4^{2,17}$

 $3.15E + 08 = N_4^{2.04}$

 $2.25E + 09 = N_4^{1,99}$

 $1.09E + 10 = N_4^{1.96}$

 $N_4 = N_1^4$

81

625

2 4 0 1

14 641

50 625

130 321

 $T_4 = p$

3

4

5

6

7

8

 N_1

3

5

7

11

15

19

р

3

4

5

6

7

8

80

Таблица б

 L_4

21 870 = $N_4^{2,27}$

 $467\ 500 = N_4^{2,03}$

 $4\ 057\ 690 = N_4^{1,95}$

48 754 530 = $N_4^{1,85}$

 $3.01E + 08 = N_4^{1.8}$

 $1,29\overline{E}+09=N_4^{1,78}$

Таблииа 5



ют схему первого измерения, а *p* схем первого измерения образуют схему второго измерения.

Аналогично, схема третьего измерения состоит из p схем второго измерения, а схема четвертого измерения состоит из p схем третьего измерения.

Сеть $\underline{C}_8 N_8$ содержит $p^4 N_4$ коммутаторов $K_4 N_4$, имеющих в совокупности $p^4 N_8$ входов. Далее вырезаются все четыре слоя мультиплексоров Mp первого каскада сети $C_8 N_8$, и входы первого слоя мультиплексоров Mp подсоединяются к входам коммутаторов $K_4 N_4$. Это возможно, так как число этих входов совпадает. Копии второго каскада сети $C_8 N_8$ в сети $\underline{C}_8 N_8$ числом p^4 разобьем на p^3 групп и обозначим через G ($1 \le G \le p^3$) номер группы и через I ($1 \le I \le p$) – номер в группе. Фактически Gзадает номер схемы первого измерения, а I – номер копии $K_4 N_4$ в схеме первого измерения.

Обозначим через J ($1 \le J \le N_4$) номер коммутатора K_4N_4 в схеме первого измерения. Обозначим через K ($1 \le K \le N_4$) номер входа каждого такого коммутатора K_4N_4 . Таким образом, вход любого коммутатора K_4N_4 задается составным индексом G, I, J, K.

Обозначим через i $(0 \le i \le p - 1)$ номер входа по mod*p* каждого мультиплексора M*p* в первом их слое. В каждом коммутаторе первого каскада сети C_8N_8 содержится p^3N_4 этих мультиплексоров M*p*. Разобьем их на N_4 групп и обозначим через *g* $(1 \le g \le N_4)$ номер группы, а через j $(1 \le j \le p)$ – номер мультиплексора M*p* в группе. Обозначим через k $(1 \le k \le N_4)$ номер коммутатора K₄N₄ в первом каскаде сети C₈N₈. Таким образом, вход любого мультиплексора M*p* в первом слое задается составным индексом *i*, *g*, *j*, *k*.

Для этого вырезаются все слои мультиплексоров М*p* первого каскада сети C_8N_8 . Первый внутренний слой содержит $W_{8,1} = N_4p^3$ мультиплексоров М*p*, входы которых соединяются с входами коммутаторов K₄N₄ в сети <u>C</u>₈N₈.

Вход мультиплексора Mp первого слоя с индексом *i*, *g*, *j*, *k* переключается на вход коммутатора K_4N_4 с индексом *G*, *I*, *J*, *K*, в котором G = i + 1, I = j, J = g и K = k. В результате образуется p^4N_4 параллельно включенных коммутаторов K_4N_4 , на входы которых поступают разреженные непересекающиеся прямые *p*-перестановки.

Вырезанные мультиплексоры Mp первого слоя объединяют выходы p^3 схем первого измерения. Вырезанные мультиплексоры Mp второго слоя объединяют выходы схем первого измерения с одинаковым индексом G, образуя выходы схем

второго уровня с этим индексом. Вырезанные мультиплексоры Mp второго слоя объединяют выходы схем второго измерения с одинаковыми индексами. Вырезанные мультиплексоры Mp третьего слоя объединяют выходы схем второго измерения с одинаковыми индексами G, образуя выходы схем третьего уровня с тем же индексом. Вырезанные мультиплексоры Mp четвертого слоя объединяют выходы схем третьего измерения, образуя выходы коммутатора K_8N_8 .

Для коммутаторов К₈N₈ справедлива

Лемма 4. Дуальный коммутатор K₈N₈ является неблокируемым коммутатором со статической маршрутизацией на любой обычной перестановке при любом р. Он обладает (σ–1)-канальной отказоустойчивостью.

Д о к а з а т е л ь с т в о. Первое утверждение опирается на использование коммутатора K_4N_4 и справедливость для него леммы 3. Второе утверждение опирается на неблокируемость коммутатора K_4N_4 и тот факт, что *p*-перестановка состоит на сечении из разреженных 1-перестановок, разведенных по разным каналам и тактам.

Канальная отказоустойчивость следует из того, что пути между источниками и приемниками в коммутаторе K_4N_4 проходят через разные схемы каждого измерения и того, что $p \ge \sigma$.

В результате образуется неблокируемый самомаршрутизируемый коммутатор K_8N_8 с $N_8 = N_4^2$ каналами, обладающий (σ – 1)-канальной отказоустойчивостью. \blacklozenge

Коммутатор K_8N_8 имеет восемь слоев выходных мультиплексоров Mp общим числом $V_8 = N_8(p^8 - 1)/(p - 1).$

Характеристики наиболее быстродействующих вариантов коммутаторов K_8N_8 приводятся в табл. 8, табл. 9 и табл. 10. В них даются характеристики коммутаторов для $\sigma = 2$ и $\sigma = 3$, которые рассчитываются по рекуррентным формулам $S_8 = N_4S_4 + p^4N_4S_4$ и $L_8 = N_4L_4 + p^4N_4L_4$. Отметим, что коммутационную и канальную сложность коммутатора K_8N_8 на основе орграфа можно сделать существенно меньше, чем у сосредоточенного коммутатора с топологией полного графа.

5. АНАЛИЗ РЕЗУЛЬТАТОВ И ВОЗМОЖНОСТИ ПРАКТИЧЕСКОГО РАЗВИТИЯ ПОСТРОЕННЫХ СЕТЕЙ

В работе предложена методика построения нового вида неблокируемых самомаршрутизируемых фотонных сетей широкой масштабируемости. Это так называемые дуальные сети, в основе которых





лежит неблокируемый дуальный коммутатор *p*×*p* с периодом разрядов в р длительностей сигнала (тактов).

Дуальный коммутатор используется как составная часть неблокируемого самомаршрутизируемого коммутатора $N_1 \times N_1$ $K_1 N_1$ с топологией квазиполного графа или орграфа. В первом случае число каналов $N_1 = p(p-1)/\sigma + 1$ и существует возможность иметь (σ-1)-канальную отказоустойчивость. Во втором случае число каналов $N_1 = p^2$ и имеется возможность увеличить число каналов. Коммутатор с топологией квазиполного (ор)графа состоит из N1 дуальных коммутаторов $p \times p$ вместе с N_1 демультиплексорами Дp 1×p и мультиплексорами Мр p×1 без линий задержки. Коммутационная сложность коммутатора К₁N₁ оценивается как $S_1 = N_1(S_0 + 2p)$. Период сигналов T_1 коммутатора $K_1 N_1$ равен периоду сигналов у дуального коммутатора, т. е. $T_1 = p$.

Из коммутаторов К₁N₁ составляются два каскада для построения блокируемой сети $N_2 \times N_2 C_2 N_2$ с $N_2 = N_1^2$ каналами. Каждый каскад состоит из N_1 коммутаторов N₁×N₁, а каналы между каскадами

 $N_8 = N_1^{8}$

6 561

390 625

5 764 801

 N_1

3

5

7

р

3

4

5

прокладываются по схеме с обменными связями. Сеть С₂N₂ преобразуется в неблокируемый самомаршрутизируемый двухкаскадный коммутатор К₂N₂ посредством ее одномерного внутреннего распараллеливания.

Если коммутатор $N_1 \times N_1$ был построен на основе квазиполного графа, то коммутатор K₂N₂ обладает (σ-1)-канальной отказоустойчивостью, поскольку $p \ge \sigma$. Коммутационная сложность коммутатора К₂N₂ оценивается по рекуррентным формулам как $S_2 = N_1 S_1 + p N_1 S_1$, а канальная сложность – как $L_2 = N_1 L_1 + p N_1 L_1$. По построению период сигналов Т₂ коммутатора К₂N₂ равен периоду сигналов коммутатора K_1N_1 , т. е. $T_2 = T_1 = p$.

Если коммутатор $N_1 \times N_1$ был построен на основе квазиполного графа, то четырехкаскадный коммутатор К₄N₄ обладает (σ-1)-канальной отказоустойчивостью, поскольку $p \ge \sigma$. Коммутационная сложность коммутатора К₄N₄ оценивается по рекуррентным формулам как $S_4 = N_2 S_2 + p^2 N_2 S_2$, а канальная сложность – как $L_4 = N_2 L_2 + p^2 N_2 L_2$. По построению период сигналов T₄ коммутатора K₄N₄ равен периоду сигналов коммутатора К₂N₂, т. е.

$$T_4 = T_2 =$$

Таблица 8

р.

Аналогично из коммутаторов К₄N₄ составляются два каскада для построения блокируемой $N_8 \times N_8$ C_8N_8 сети с $N_8 = N_4^2 = N_1^8$ каналами.

Каждый каскад состоит из N₄ коммутаторов К₄N₄, а каналы между каскадами прокладываются по схеме с обменными связями.

Коммутационная

формулам

сложность коммутатора К₈N₈ оценивается по ре-

Сеть С₈N₈ преобразуется в неблокируемый самомаршрутизируемый двухкаскадный коммутатор $K_8 N_8$ посредством ее 145 260 540 = $N_8^{2,14}$ четырехмерного BHVT- $7.509E + 10 = N_8^{1.94}$ реннего распараллелива- $6,099E + 12 = N_8^{1,89}$ ния. Если коммутатор $N_1 \times N_1$ был построен на основе квазиполного Таблица 10 графа, то и восьмикаскадный коммутатор K_8N_8 обладает (σ-1)-канальной отказоустойчиво-

стью.

Характеристики дуальных коммутаторов К₈ N₈ на основе орграфа

р	N_1	$N_8 = N_1^{\circ}$	$T_8 = p$	S_8	L_8
2	4	65 536	2	$189\ 399\ 040 = N_8^{1,72}$	77 987 840 = $N_8^{1,64}$
3	9	5 764 801	3	$3,283\mathrm{E} + 12 = N_8^{1,64}$	$9,531\mathrm{E} + 11 = N_8^{1,57}$
4	16	4,29E+09	4	$3,678\mathrm{E}+15=N_8^{1,62}$	$8,256\mathrm{E}+14=N_8^{1,55}$

Характеристики дуальных коммутаторов К₈ N₈ с одноканальной отказоустойчивостью

р	N_1	$N_8 = N_1^{8}$	$T_8 = p$	S_8	L_8
2	2	256	2	739 840 = $N_8^{2,44}$	$304\ 640 = N_8^{2,28}$
3	4	65 536	3	$4,998\mathrm{E} + 09 = N_8^{2,01}$	$1,451\mathrm{E} + 09 = N_8^{1,9}$
4	7	5 764 801	4	$4,937\mathrm{E} + 12 = N_8^{1,88}$	$1,108\mathrm{E}+12=N_8^{1,78}$

Характеристики дуальных коммутаторов K₈ N₈ с двухканальной

отказоустойчивостью

 S_8

 $\overline{500\ 341\ 860} = N_8^{2,28}$

 $3,345E + 11 = N_8^{2,06}$

 $3,331E + 13 = N_8^2$

 $T_8 = p$

3

4

5

Таблица 9

 L_8

CONTROL SCIENCES No.5 • 2021

куррентным

как $S_8 = N_4 S_4 + p^4 N_4 S_4$, а канальная сложность – как $L_8 = N_4 L_4 + p^4 N_4 L_4$. По построению период сигналов T_8 коммутатора K_8N_8 равен периоду сигналов коммутатора K_4N_4 , т. е. $T_8 = T_4 = p$.

Характеристики коммутаторов К₂N₂, К₄N₄ и К₈N₈ обладают несколькими степенями свободы. Прежде всего, при увеличении основания р число каналов увеличивается, а быстродействие уменьшается. Кроме того, показательная сложность уменьшается при увеличении основания р и имеется возможность разменивать быстродействие на сложность. К тому же, увеличение числа каналов посредством увеличения числа каскадов также сопровождается уменьшением показательной сложности.

Предложенная методика позволяет строить неблокируемые самомаршрутизируемые сети с самоподобной структурой. Коммутатор K₂N₂ состоит из дуальных коммутаторов K₁N₁ с топологией дуального графа или дуального орграфа и использует одномерное внутреннее распараллеливание. Коммутатор K_4N_4 состоит из коммутаторов K_2N_2 и использует двумерное внутреннее распараллеливание. Наконец, коммутатор K_8N_8 состоит из коммутаторов К₄N₄ и использует четырехмерное внутреннее распараллеливание. Все эти коммутаторы наследуют базовые свойства коммутатора K₁N₁, такие как неблокируемость при статической самомаршрутизации и канальную отказоустойчи-

вость (при необходимости), но при существенно меньшей сложности.

Широкая масштабируемость неблокируемых коммутаторов также может быть достигнута многократным применением метода инвариантного расширения к коммутатору К₁N₁ с топологией орграфа, в котором используется обычный коммутатор $p \times p$. Такие расширенные коммутаторы имеют период разрядов в один такт, но обладают повышенной сложностью. В табл. 11 сравнивается коммутационная сложность дуальных коммутаторов K_4N_4 и K_8N_8 и расширенных коммутаторов K_1N_1 . Видно, что дуальные коммутаторы имеют на несколько порядков меньшую коммутационную сложность.

Заметим, что дуальный коммутатор разрешает конфликты шинным способом только в первом каскаде в коммутаторе K_1N_1 . Все остальные конфликты во всех каскадах просто предотвращаются благодаря использованию внутреннего распараллеливания, и в них дуальные коммутаторы используются как обычные коммутаторы $p \times p$. Поэтому, может быть, имеет смысл использовать дуальный коммутатор в его первоначальном виде [1-3] – в виде пары «мультиплексор – демультиплексор», коммутационная сложность которой в p раз меньше. Это даст уменьшение коммутационной сложности дуальных коммутаторов К₄N₄ и К₈N₈ еще в несколько раз (1,5÷4,5).

Таблица 11

Ко	Коммутационные сложности неблокируемых четырехкаскадных коммутаторов ($S_{4, D}$)							
	и расширенн	ных коммутаторов с тополе	огией квазиполных орграфо	в (<i>S</i> _{PO})				
р	N_4	Дуальный коммутатор Расширенный коммута- Отношение						
		K_4N_4	тор $K_1 N_1$	$S_{ m PO}/S_{4, m Z}$				
		S _{4, Д}	$S_{ m PO}$					
2	256	$46\ 080 = N_4^{1,94}$	$261\ 120 = N_4^{2,25}$	5,67				
3	6 561	$6\ 298\ 560 = N_4^{1,78}$	129 120 480 = $N_4^{2,12}$	20,5				
4	65 536	22 282 400 = $N_4^{1,74}$	11 453 071 360 = $N_4^{2,09}$	514				
Ко	оммутационные	сложности неблокируемы	х восьмикаскадных коммута	торов (S _{8, Д})				
	и расширенн	ных коммутаторов с тополе	огией квазиполных орграфо	в (S _{PO})				
р	N_8	Дуальный коммутатор	Расширенный коммута-	Отношение				
		K_8N_8	тор $K_1 N_1$	$S_{ m PO}/S_{8, m Z}$				
		S _{8, Д}	$S_{ m PO}$					
2	65 536	$18\ 939\ 9040 = N_8^{1,72}$	17 179 607 040 = $N_8^{2,12}$	85,7				
3	43 046 721	$3,283\mathrm{E}+12=N_8^{1,64}$	$5,55822E + 15 = N_8^{2,06}$	1 640				
4	4,29E + 09	$3,678E + 15 = N_8^{1,62}$	$4,91906E+19 = N_8^{2,04}$	13 107				

Сравнительная сложность дуальных коммутаторов и расширенных орграфов

Отметим, что сложотказоустойчиность вых коммутаторов K_2N_2 и К₄N₄ при малых *р* больше сложности полного графа, а при больших *р* – меньше. При этом сложность коммутаторов K₈N₈ оказывается существенно меньше сложности полного графа при любых р. Специально хотелось бы выделить характеристики коммутатора K_8N_8 при p = 2 и $\sigma = 1$. При числе каналов $N_8 = 65536$ И влвое меньшем быстродействии его коммутационная сложность сопоставима со сложностью пятикаскадной небло-



кируемой сети Клоза на базе 64-канального маршрутизатора YARC [9] с числом каналов N = 32768, при условии, что она построена как неблокируемая сеть [7, 8]. Оценка сложности этой неблокируемой сети Клоза дает величину $S = N^{1,73}$. Однако эта сеть не имеет параллельных процедур ни статической, ни динамической самомаршрутизации.

Остальные коммутаторы K_8N_8 с p > 2 и $\sigma = 1$ имеют еще меньшую коммутационную сложность и еще большую масштабируемость, но при меньшем быстродействии.

Пониженное в р раз быстродействие коммутаторов K_2N_2 , K_4N_4 и K_8N_8 может быть скомпенсировано разными протокольными способами. Можно использовать процессоры с р независимыми портами, делить пакеты на р частей и передавать их параллельно. Широкая масштабируемость построенных коммутаторов легко обеспечивает такой режим работы, но посредством уменьшения в *p* раз числа абонентов и увеличения сложности сети. Также можно использовать параллельнопоследовательный способ передачи пакетов по р линиям как в протоколе PCI-express уже без уменьшения числа абонентов.

Одним из недостатков коммутаторов K_1N_1 , K_2N_2 , K_4N_4 и K_8N_8 является их оптимизация под бесконфликтную реализацию произвольных перестановок. А каким будет их поведение на произвольном трафике? Для его определения можно придать мультиплексорам в выходных каскадах свойство «одноканальности». Мультиплексор с таким свойством из нескольких входных пакетов пропускает на выход только один и блокирует прохождение остальных пакетов. Передача блокированных пакетов, не подтвержденных приемниками, повторно осуществляется источниками.

Существенным недостатком предложенной методики является необходимость параллельной передачи сигнальной и управляющей информации, что существенно увеличивает необходимую полосу пропускания. Для фотонных коммутаторов этот недостаток не является роковым, так как оптический кабель может нести одновременно сотни разных частот. Однако в общем случае этот недостаток можно ликвидировать, если обеспечить поразрядную синхронизацию сигналов разных каналов. Это можно сделать на основе метода [34, 35] с локацией взаимного положения источников и приемников и соответствующих задержек передач от источников. В этом случае управляющую информацию для дуальных коммутаторов и демультиплексоров можно передавать, как обычно, в виде наборов разрядов в заголовке пакетов.

Приятным бонусом поразрядной синхронизации является возможность построения в канале на входе каждого приемника АЛУ только сетевыми средствами. Такие АЛУ были разработаны для метода вычислений в общем канале [36]. Для их реализации требуется передача по каналу значений разрядов чисел в парафазном виде – по двум линиям с активными сигналами для значений 0 и 1 в каждом. В сетевом АЛУ производится операция над числом, поступающим по каналу, и числом, находящимся у приемника, с образованием результата в канале после АЛУ. В канале возможно выполнение операций сложения, умножения, любых поразрядных логических операций, в том числе нахождения максимума (минимума).

ЗАКЛЮЧЕНИЕ

В работе предложена методика построения неблокируемых отказоустойчивых фотонных сетей широкой масштабируемости, рассмотренных в статье [5], но со значительной меньшей сложностью. Эта методика базируется на трех основных компонентах:

• использование *p*-канального дуального коммутатора с периодом разрядов в *p* тактов, который оказывается неблокируемым на любом входном трафике — именно это является необходимым условием неблокируемости более сложных сетей;

• использование коммутатора с топологией квазиполного графа или орграфа с дуальным коммутатором внутри, в результате чего сохраняется неблокируемость, обеспечивается канальная отказоустойчивость и более высокая масштабируемость при каскадировании по сравнению с чистым дуальным коммутатором;

• применение метода внутреннего распараллеливания для сохранения неблокируемости посредством предотвращения конфликтов и сохранения отказоустойчивости, которое обеспечивает широкую масштабируемость при каскадировании неблокируемых сетей.

В работе [5] масштабирование осуществляется путем каскадного применения метода инвариантного расширения посредством использования дополнительных внешних мультиплексоров и демультиплексоров. В настоящей работе масштабирование осуществляется путем каскадирования неблокируемых сетей меньшего размера и применения обобщенного метода внутреннего распараллеливания на каждом шаге каскадирования.

Каскадирование неблокируемой сети с *N* каналами осуществляется посредством построения



блокируемой сети с N^2 каналами. Эта сеть состоит из двух каскадов с обменными связями с N исходными неблокируемыми сетями в каждом. Блокировки в этой двухкаскадной сети возникают на выходных мультиплексорах первого каскада. Эти блокировки предотвращаются посредством разведения конфликтующих каналов по нескольким копиям второго каскада и перемещения мультиплексоров на выходы той частью второго каскада, которая выполняет маршрутизацию пакетов. В этой части сети конфликты не возникают, так как она состоит из копий неблокируемых подсетей, которые выполняют маршрутизации разреженных перестановок. Объединение разреженных перестановок в полную перестановку на сети с N^2 каналами осуществляется перемещенными каскадами мультиплексоров бесконфликтно.

При первом каскадировании [5] внутреннее распараллеливание осуществляется с использованием р копий второго каскада и использованием однослойного каскада выходных мультиплексоров. Второе каскадирование при построении неблокируемой сети с N^4 каналами выполняется с использованием p² копий второго каскада и использованием двухслойного каскада выходных мультиплексоров. Третье каскадирование при построении неблокируемой сети с N⁸ каналами выполняется с использованием p⁴ копий второго каскада и использованием четырехслойного каскада выходных мультиплексоров. Таким образом, были построены неблокируемые двух-, четырех- и восьмикаскадные сети с каскадами, состоящими из неблокируемых дуальных сетей с топологией квазиполного графа или орграфа.

Внутреннее распараллеливание при каждом каскадировании сохраняет период разрядов и сопровождается уменьшением удельной сложности создаваемой неблокируемой сети. В частности, были построены неблокируемые сети, имеющие удельную сложность не больше, чем у теоретической неблокируемой сети Клоза.

Рассмотренная методика может служить основой для построения практических проектов неблокируемых коммутаторов широкой масштабируемости со статической самомаршрутизацией и канальной отказоустойчивостью.

ЛИТЕРАТУРА

 Барабанова Е.А., Вытовтов К.А., Подлазов В.С. Многокаскадные коммутаторы для оптических и электронных суперкомпьютерных систем // Материалы 8-го Национального Суперкомпьютерного Форума (НСКФ-2019). – Переславль-Залесский, 2019. – http://2019.nscf.ru/TesisAll/02_ Apparatura/037_BarabanovaEA.pdf. [Barabanova, E.A., Vytovtov, K.A., Podlazov, V.S. Mnogokaskadnye kommutatory dlya opticheskih i elektronnyh superkomp'yuternyh sistem // Materialy 8-go Nacional'nogo Superkomp'yuternogo Foruma (NSKF-2019). – Pereslavl'-Zalesskij, 2019. – http:// 2019.nscf.ru/TesisAll/02_Apparatura/037_BarabanovaEA.pdf. (In Russian)]

- 2. Барабанова Е.А., Вытовтов К.А., Вишневский В.М., Подлазов В.С. Новый принцип построения оптических устройств обработки информации для информационно-измерительных систем // Датчики и системы. – 2019. – № 9. – С. 3–9. [Barabanova, E.A., Vytovtov, K.A., Vishnevskij, .M., Podlazov, V.S. The New Principle for the Construction of Optical Information Processing Devices for Information-Measuring Systems // Sensors and Systems. – 2019. – No. 9. – P. 3–9 (In Russian)]
- Barabanova, E., Vytovtov, K., Vishnevskiy, V., Podlazov, V. Model of Optical Non-blocking Information Processing System for Next-generation Telecommunication Networks // Proceedings of the 22nd International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications (DCCN-2019). – Moscow, 2019. – P. 188– 198. – DOI: 10.1007/978-3-030-36625-4_16.
- 4. 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.]
- Барабанова Е.А., Вытовтов К.А., Подлазов В.С. Двухкаскадные дуальные фотонные коммутаторы в расширенном схемном базисе // Проблемы управления. – 2021. – № 1. – С. 69–80. [Barabanova, E.A., Vytovtov, K.A., Podlazov, V.S. Two-Stage Dual Photonic Switches in an Extended Circuit Basis // Control Sciences. – 2021. – No. 1. – Р. 69–80.]
- Барабанова Е.А., Вытовтов К.А., Подлазов В.С. Неблокируемые отказоустойчивые двухкаскадные дуальные фотонные коммутаторы // Проблемы управления. – 2021. – № 4. – С. 82–92. [Barabanova, E.A., Vytovtov, K.A., Podlazov, V.S. Non-blocking Fault-Tolerant Two-Stage Dual Photon Switches. – Control Sciences. – 2021. – No. 4. – Р. 67–76.]
- Clos, C. A Study of Non-locking Switching Networks // Bell System Tech. J. – 1953. – Vol. 32. – P. 406–424.
- Бенеш В.Э. Математические основы теории телефонных сообщений. – М.: Связь. – 1968. – С. 83–150. [Benesh, V.E. Matematicheskie osnovy teorii telefonnyh soobshhenij. – М.: Svjaz'. – 1968. – Р. 83–150. (In Russian)]
- Scott, S., Abts, D., Kim, J., and Dally, W. The Black Widow High-radix Clos Network // Proc. 33rd Intern. Symp. Comp. Arch. (ISCA'2006). – Boston, 2006. – URL: https://www.researchgate.net/publication/4244660_The_Black Widow_High-Radix_Clos_Network.
- 10.De Sensi, D., Di Girolamo, S., H. McMahon, K., Roweth, D. An In-Depth Analysis of the Slingshot Interconnect // arXiv: 2008.08886v1. – 2020. – URL: https://www.researchgate.net/ publication/343786515_An_In-Depth_Analysis_of_the_ Slingshot_Interconnect.
- 11.Alverson, R., Roweth, D., and Kaplan, L. The Gemini System Interconnect // 18th IEEE Symposium on High Performance Interconnects. – Santa Clara, CA, 2009. – P. 83–87.
- 12.Alverson, R., Roweth, D., Kaplan, L., and Roweth, D. Cray XC[®] Series Network. URL: http://www.cray.com/ Assets/PDF/products/xc/CrayXC30Networking.pdf.
- 13.Kim, J., Dally, W. J., Scott, S. and Abts, D. Technology-Driven, Highly-Scalable Dragonfly Topology // Proceedings of the 35th Annual International Symposium on Computer Architecture (ISCA'2008). – Beijing, 2008. – P. 77–88. – URL:





http://users.ece.gatech.edu/~sudha/academic/class/Networks/Le ctures/4%20-%20Topologies/papers/dragonfly.pdf.

- 14.Mellanox OFED for Linux User Manual. Rev 2.3-1.0.1. Melanox Technologies, Ltd.: 2014. – URL: https://dlcdnets. asus.com/pub/ASUS/mb/accessory/PEM-FDR/Manual/Mellanox_OFED_Linux_User_Manual_v2_3-1_0_1.pdf.
- 15.Pipenger, N. On Rearrangeable and Non-blocking Switching Networks // J. Comput. Syst. Sci. – 1978. – Vol. 17. – P. 307– 311.
- 16.*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.
- 17. *Tzeng, N., Wei, S.* Enhanced Hypercubes // IEEE Trans. Computers. 1991. Vol. 40, no. 3. P. 284–294.
- 18.*Efe*, K. A Variation on the Hypercube with Lower Diameter // IEEE Trans. Computers. – 1991. – Vol. 40, no. 11. – P. 1312– 1316.
- 19.Kim, J., Dally, W.J. Flattened Butterfly Topology for On-Chip Networks // IEEE Computer Architecture Letters. – 2007. – Vol. 6, no. 2. – P. 37–40.
- 20.*Gu, Q.P. and Tamaki, H.* Routing a Permutation in Hypercube by Two Sets of Edge-Disjoint Paths // J. of Parallel and Distributed Comput. 1997. Vol. 44, no. 2. P. 147–152.
- 21.*Lubiw*, A. Counterexample to a Conjecture of Szymanski on Hypercube Routing // Inform. Proc. Let. 1990. – Vol. 35(2). – P. 57–61.
- 22.*Stepanenko, S.* Structure and Implementation Principles of a Photonic Computer // EPJ Web of Conferences. 2019. Vol. 224. DOI: https://doi.org/10.1051/epjconf/201922404002.
- 23.Жабин И.А., Макагон Д.В., Поляков Д.А. и др. Первое поколение высокоскоростной коммуникационной сети «Ангара» // Наукоемкие технологии. – 2014. – № 1. – С. 21–27. [Zhabin, I.A., Makagon, D.V., Polyakov, D.A., et al. Pervoe pokolenie vysokoskorostnoi kommunikatsionnoi seti «Angara» // Naukoemkie tekhnologii. – 2014. – No. 1. – Р. 21–27. (In Russian)]
- 24.Stegailov, V., Agarkov, A., Biryukov, V., et al. Early Performance Evaluation of the Hybrid Cluster with Torus Interconnect Aimed at Molecular Dynamics Simulations // International Conference on Parallel Processing and Applied Mathematics. Cham: Springer, 2017. P. 327–336.
- 25.Ajima, Y., Inoue, T., Hiramoto, S., Shimiz, T. Tofu: Interconnect for the K Computer // Fujitsu Scientific & Technical Journal. – Vol. 48, no. 3. – July 2012. – P. 280–285. – URL: https://www.researchgate.net/publication/265227674_Tofu_Int erconnect_for_the_K_computer.
- 26.Arimili, B., Arimilli, A., Chung, V., et al. The PERCS High-Performance Interconnect // 18th IEEE Symposium on High Performance Interconnects. – New York, 2009. – P. 75–82.
- 27.Kathareios, G., Minkenberg, C., Prisacari, B., et al. Cost-Effective Diameter-Two Topologies: Analysis and Evaluation // SC'15: Proceedings of the International Conference for High Performance Computing. Networking, Storage and Analysis (SC'15). – 2015. – P. 1–11.
- 28.Besta, M. and Hoefler, T. Slim fly: A Cost Effective Low-Diameter Network Topology // SC'14: Proceedings of the In-

ternational Conference for High Performance Computing. Networking, Storage and Analysis. – 2014. – P. 348–359

- 29.Flajslik, M., Borch, E., and Parker, M.A. Megafly: A topology for exascale systems // Yokota, R., Weiland, M., Keyes, D., and Trinitis, C. eds., High Performance Computing. Cham. Springer International Publishing. – 2018. – P. 289–310.
- 30.Ahn, J. H., Binkert, N., Davis, A., et al. Hyperx: Topology, Routing, and Packaging of Efficient Large-Scale Networks // Proceedings of the Conference on High Performance Computing Networking. Storage and Analysis. – 2009. – P. 1–11.
- 31.Domke, J., Matsuoka, S., Ivanov, I.R., et al. Hyperx Topology: First At-scale Implementation and Comparison to the Fat-Tree // Proceedings of the International Conference for High Performance Computing, Networking. Storage and Analysis (SC'19). New York, USA, Association for Computing Machinery. – 2019.
- 32.Singla, A., Hong, C.-Y., Popa, L., and Godfrey, P.B. Jellyfish: Networking Data Centers Randomly // Presented as part of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI 12). San Jose. CA. USENIX. – 2012. – P. 225–238.
- 33.Холл М. Комбинаторика. Главы 10-12. М. Мир. 1970. 424 с. [Hall, M. Combinatorial Theory. – Blaisdell Publishing Company. Waltham. – 1967.]
- 34.*Стецюра Г.Г.* Компьютерная сеть с быстрой распределенной перестройкой своей структуры и обработкой данных в процессе их передачи // Проблемы управления. 2017. № 1. С. 47–56. http://pu.mtas.ru/archive/Stetsyura_117.pdf [*Stetsyura, G.G.* Computer Network with the Fast Distributed Reorganization of Its Structure and Data Processing During Their Transmission // Control Sciences. 2017. No. 1. P. 47–56.] URL: (In Russian)]
- 35.*Стецюра Г.Г.* Компьютерные кластеры с быстрым аппаратным выполнением синхронизации сообщений и распределенных вычислений сетевыми средствами // Проблемы управления. 2020. № 4. С. 61–69. [*Stetsyura, G.G.* The Computer Clusters with Fast Synchronization of Messages and with Fast Distributed Computing by the Network Hardware // Control Sciences. 2020. No. 4. Р. 61–69. (In Russian)]
- 36.Прангишвили И.В., Подлазов В.С., Стецюра Г.Г. Локальные микропроцессорные вычислительные сети. Глава шестая. М.: Наука, 1984. 175 с. [Prangishvili, I.V., Podlazov, V.S., Stetsyura, G.G. Lokal'nye mikroprocessornye vychislitel'nye seti. Glava shestaya. М.: Nauka, 1984. 175 р. (In Russian)]

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

Поступила в редакцию 25.03.2021, после доработки 12.08.2021. Принята к публикации 24.08.2021

Подлазов Виктор Сергеевич – д-р техн. наук, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, podlazov@ipu.ru,



NON-BLOCKING FAULT-TOLERANT DUAL PHOTON SWITCHES WITH HIGH SCALABILITY

V.S. Podlazov

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

⊠ podlazov@ipu.ru

Abstract. This paper continues the construction of a fundamentally new class of system area networks (dual photon networks) with the following features: non-blocking property and static self-routing, high scalability with the maximum achievable speed and a small complexity compared to a full switch, and balancing the scalability-speed and complexity-speed ratios. These networks are implemented in an extended circuit basis consisting of dual photon switches and separate photon multiplexers and demultiplexers. We propose a method for constructing a fault-tolerant dual network with the indicated properties based on networks with the quasi-complete graph and quasi-complete digraph topologies and the invariant extension method with internal parallelization. Also, we propose a method for extending the two-stage dual network designed previously into four-stage and eight-stage dual networks with high scalability while maintaining the original network period and reducing its exponential complexity.

Keywords: photon switch, dual switch, photon multiplexers and demultiplexers, multistage switch, conflict-free self-routing, non-blocking switch, static self-routing, quasi-complete digraph, quasi-complete graph, invariant extension of networks, switching properties, direct channels, scalability and speed.