

РАСШИРЕНИЕ ВОЗМОЖНОСТЕЙ СИСТЕМНОЙ СЕТИ «АНГАРА»

В.С. Подлазов, М.Ф. Каравай

Рассмотрен способ расширения возможностей системной сети «Ангара» благодаря совместному применению собственных 24-портовых маршрутизаторов и рыночных «хабов» 1×3 и 1×4 . Способ основан на применении сетей с топологией квазиполных графов, построенных в расширенном элементном базисе, состоящем из коммутаторов (маршрутизаторов) и хабов дуплексных каналов. Эти сети лежат в основе метода инвариантного расширения произвольных сетей, сохраняющего задержки передачи расширяемых сетей и наращивающего число их абонентов. Этим методом из маршрутизаторов сети «Ангара» построен распределенный расширенный маршрутизатор, и на его основе построены сети со структурой 1, 2, 3, 4-мерных торов либо содержащие существенно большее число абонентов при одинаковом диаметре сети, либо имеющие меньший диаметр при одинаковом числе абонентов по сравнению с сетью «Ангара». Фактически рассмотренный способ дал возможность увеличить число процессоров, сократить диаметр и повысить канальную отказоустойчивость сети.

Ключевые слова: системные сети суперкомпьютеров, сеть «Ангара», сети с топологией квазиполных графов, инвариантное расширение сетей, число абонентов сети и диаметр сети, масштабирование и быстродействие.

ВВЕДЕНИЕ

В настоящее время используется небольшое число базовых структур системных сетей суперкомпьютеров — сложенная сеть Клоза, толстое дерево, многомерный тор, обобщенный гиперкуб и двухуровневая склейка полных графов. Среди них нет идеальных — выбор той или иной структуры сразу задает и ряд ее ограничений-недостатков. Так, сети с топологией D -мерных торов наименее сложные по схемным затратам, но обладают наибольшими задержками передачи данных. Поэтому разработка методов устранения недостатков выбранной структуры в рамках ее базовых возможностей является актуальной задачей совершенствования современных системных сетей. Она представляет собой задачу проектного управления характеристиками сети с целью улучшения (даже оптимизации) функционально важных ее характеристик.

Системные сети современных суперкомпьютеров строятся на базе многопортового маршрутизатора — единого однокристалльного 48-портового

маршрутизатора *YARC* для 3-мерного тора *Gemini* и для 4-мерного обобщенного гиперкуба *Dragonfly* фирмы CRAY [1, 2]. При этом наблюдается тенденция перехода на топологию гиперкуба как сети с меньшим диаметром и большим быстродействием.

В России в настоящее время нет таких маршрутизаторов. Имеется функционально полный однокристалльный маршрутизатор сети «Ангара» с 8-ю дуплексными портами [3]. Он предназначен для построения системной сети в виде 4-мерного тора с числом узлов $R = 4K$ и диаметром $D = 16$ скачков. Потом было заявлено о возможности иметь на нем $R = 16 \div 32K$ процессоров. Если это увеличение достигается путем удвоения числа узлов в измерениях, то оно понижает быстродействие сети как из-за увеличения диаметра до $D = 32$, так и из-за увеличения задержек в очередях к кольцам, так как отдельное кольцо параллельно пропускает при равномерном трафике в среднем не более 8 пакетов [4]. В других торах [1] последняя проблема снимается благодаря применению параллельных колец в каждом измерении.

Таблица 1

**Характеристики сети «Ангара» в формате R/D
(число узлов/диаметр)**

Число маршрутизаторов в измерении	2	8	16
1-мерная решетка	40/6	128/11	256/19
1-мерный тор	32/6	128/8	256/12
2-мерный тор	32/6	502/13	2048/21

В работе [5] было предложено создавать сборные маршрутизаторы с большим числом портов путем сцепления нескольких (двух) маршрутизаторов. Для сети «Ангара» этот подход был реализован при создании однокорпусного 24-портового маршрутизатора [6–8] путем сцепления четырех 8-портовых маршрутизаторов. Применение такого маршрутизатора резко упрощает построение сетей самых разных размеров — от десятков процессоров до нескольких их тысяч (в топологии 1-мерной решетки, 1-мерного или 2-мерного тора). В таком виде сеть «Ангара» считается базовой сетью для построения отечественных суперкомпьютеров.

Заметим, что 24-портовый маршрутизатор имеет внутренний диаметр в 4 скачка: 1 скачок от входного порта до соединительного порта в 8-портовом маршрутизаторе, 2 скачка между 8-портовыми маршрутизаторами и 1 скачок от соединительного порта до выходного порта. Однако в сети для связи между 24-портовыми маршрутизаторами применяются 4 дуплексных канала: между заданными 8-портовыми маршрутизаторами. Это делает проходную задержку равной 1 скачку.

В табл. 1 представлены типовые структуры сети «Ангара» и их характеристики. Видно, что решетку разумно применять только в минимальном варианте, а торы, наоборот, в вариантах со многими маршрутизаторами. Заметим, что 3-мерный тор на базе 24-портового коммутатора уже не может быть создан из-за недостаточного числа портов, что делает невозможным дальнейшее увеличение числа процессоров в сети «Ангара» без увеличения числа узлов в кольцах и задержек передачи по ним.

Спектр характеристик сети «Ангара» можно существенно расширить, если воспользоваться внешней сцепкой 24-портовых маршрутизаторов с «хабами» 1×3 ; 8-портовый маршрутизатор одновременно является и «хабом» 1×3 . Однако проще применять «хабы» в виде имеющихся на рынке разветвителей 1×3 или 1×4 дуплексных каналов интерфейса *PCI-express*. В идеальном случае «хаб» размещается на материнской плате процессора, т. е. применяется абонент сети с внутренним «хабом».

В настоящей статье рассматриваются два подхода к расширению возможностей сети «Ангара»: подход снизу — путем расширения 24-портового маршрутизатора и подход сверху — путем расширения исходной сети. При этом рассматриваются разные критерии формирования структуры сети — по минимальному диаметру, по максимальному числу абонентов и по сложности сети, выраженной в числе 24-портовых маршрутизаторов. Считается, что «хаб» имеет диаметр в 1 скачок — на пути к процессору.

**1. ИНВАРИАНТНОЕ РАСШИРЕНИЕ
ПРОИЗВОЛЬНЫХ СЕТЕЙ**

Рассмотрим задачу инвариантного расширения произвольной системной сети (СС) [9]. Пусть имеется исходная СС (рис. 1), объединяющая K абонентов — $ИсхС(K)$. Необходимо построить расширенную сеть $РасС(R)$, объединяющую $R > K$ абонентов, которая состоит из копий сетей $ИсхС(K)$, имеющих непересекающиеся наборы абонентов, и сохраняет ее маршрутные свойства.

Например, сеть (см. рис. 1) можно расширить, увеличив число абонентов втрое и число портов каждого абонента вдвое и сохранив как инвариант возможность соединения между любыми двумя абонентами только через одну сеть $ИсхС(K)$. Для этого потребуются три копии сети $ИсхС(K)$ и каждого абонента подсоединить к двум разным копиям сети $ИсхС(K)$. Конкретные схемы соединений при заданных числе портов абонента m и исходной сети K рассматриваются далее. Заметим, что маршрутные свойства расширенной сети остались такими же, как в исходной сети.

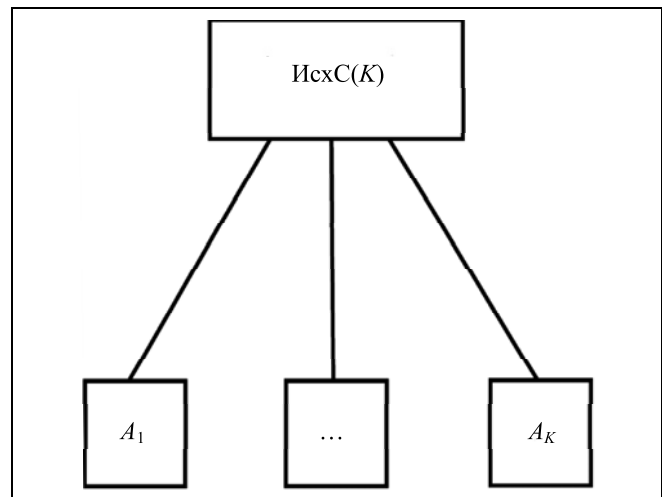


Рис. 1. Исходная произвольная сеть на K абонентов с одним портом у каждого абонента



В данной работе маршрутные свойства — это способы маршрутизации пакетов данных, которые обеспечивают минимальные их задержки при передаче (минимальный диаметр) и максимальное число абонентов сети (процессоров).

Решение задачи расширения сети ищется путем увеличения числа портов у абонентов до $m > 1$ и применения нескольких копий сети $ИсхС(K)$, имеющих разные наборы абонентов. Конкретно, решением является расширенная сеть $РасС(R)$, имеющая структуру двудольного графа, в котором все $N(m)$ узлов одной доли имеют одинаковые степени K , а все R узлов другой доли — одинаковые степени m . Здесь функция $N(m)$ еще не задана, и она определяется далее. Ребра между узлами разных долей проводятся так, что между любыми двумя узлами одной доли имеются только пути длины 2, каждый из которых проходит через один узел другой доли. Таких путей должно быть не менее $\sigma \geq 1$, и все они должны проходить через разные узлы другой доли.

В расширенной сети $РасС(R)$ копии $ИсхС(K)$ трактуются как вершины одной доли двудольного графа, абоненты с m дуплексными портами — как вершины другой его доли, а дуплексные каналы между ними — как его ребра. Иначе говоря, $РасС(R)$ — это сеть, к которой подсоединено R абонентов и которая состоит из N копий $ИсхС(K)$, к каждой из которых подсоединено точно K разных абонентов, и каждый абонент подсоединен к m копиям $ИсхС(K)$, и каждый абонент соединяется с любым другим абонентом не менее чем через $\sigma \geq 1$ разных копий $ИсхС(K)$.

Из приведенного описания следует, что в сети $РасС(R)$ любые два абонента соединены друг с другом через одну сеть $ИсхС(K)$. Это означает сохранение в сети $РасС(R)$ маршрутных свойств сети $ИсхС(K)$, которые оказываются инвариантом рассмотренного метода расширения. В частности, задержки доставки пакетов данных в сети $РасС(R)$ оказываются практически неизменными. Кроме того, такое расширение может при $\sigma > 1$ обеспечить и канальную отказоустойчивость сети $РасС(R)$ или, наоборот, повысить ее удельную (на один процессор) пропускную способность.

В простейшем двудольном графе выбирается $K = m$ и $R = N(m)$. Если при этом значение m выбирается минимальным, при котором между любыми двумя вершинами одной доли существует точно σ путей длины 2, то такой двудольный граф мы называем минимальным квазиполным графом.

Здесь возникает два вопроса: существуют ли квазиполные графы и как их построить? Ответ на первый дает комбинаторика, которая давно изучает такие объекты, как симметричные блок-схемы,

изоморфные минимальным квазиполным графам [9, 10]. К сожалению, полного ответа на второй вопрос комбинаторика не дает, и в работе [11] рассматривается ряд способов построения некоторых симметричных блок-схем.

Необходимое условие существования квазиполных графов задается выражением $N = m(m - 1)/\sigma + 1$ [10].

В математической теории блок-схемы строятся как оптимальные объекты с максимальным N при заданных m и σ и минимальным числом связей между ее частями. Поэтому все расширения исходной сети с помощью минимальных квазиполных графов являются оптимальными по числу абонентов и по уровню задержек в передачах между абонентами. Правда, эта оптимизация достигается путем увеличения сложности сети. Это ожидаемый результат, так как из трех противоречивых параметров сети (числа узлов, задержек передачи и сложности сети) можно оптимизировать только два.

Любой квазиполный граф задается таблицей межсоединений, состоящей из N строк и $m + 1$ столбцов. Первый столбец содержит номера копий $ИсхС(m)$, занумерованных от 1 до N . В m ячейках i -й строки, $1 \leq i \leq N$, содержатся номера абонентов, соединенных дуплексным каналом с i -й копией $ИсхС(m)$. Табл. 2 задает квазиполный граф при $m = 3$ и $\sigma = 1$.

Сеть $РасС(N)$ со схемой минимального квазиполного графа называется простейшей СС и обозначается как ПРС(N, m, σ). Например, табл. 1 задает также межсоединения в ПРС(7, 3, 1). На рис. 2 изображена схема ПРС(7, 3, 1) в виде минимального квазиполного графа.

Табл. 3 задает ПРС(7, 4, 2) с $\sigma = 2$ разными путями между любыми двумя абонентами. На рис. 3 изображена схема этой расширенной СС в виде минимального квазиполного графа, жирными ребрами выделены разные пути между двумя парами абонентов, отмеченных разными заливками.

Задача построения расширенной сети $РасС(R)$ решается следующим образом. Увеличивается чис-

Таблица 2

Межсоединения в квазиполном графе при $m = 3$ и $\sigma = 1$

Копии ИсхС(3)	Абоненты		
	1	1	7
2	2	1	6
3	3	2	7
4	4	3	1
5	5	4	2
6	6	5	3
7	7	6	4

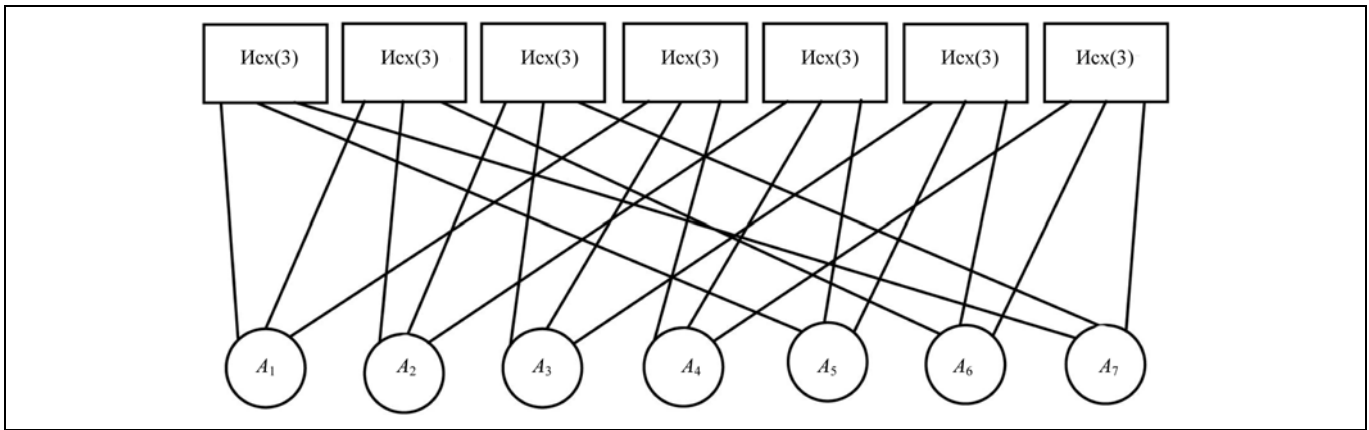


Рис. 2. Схема ПРС(7, 3, 1) в виде минимального квазиполного графа

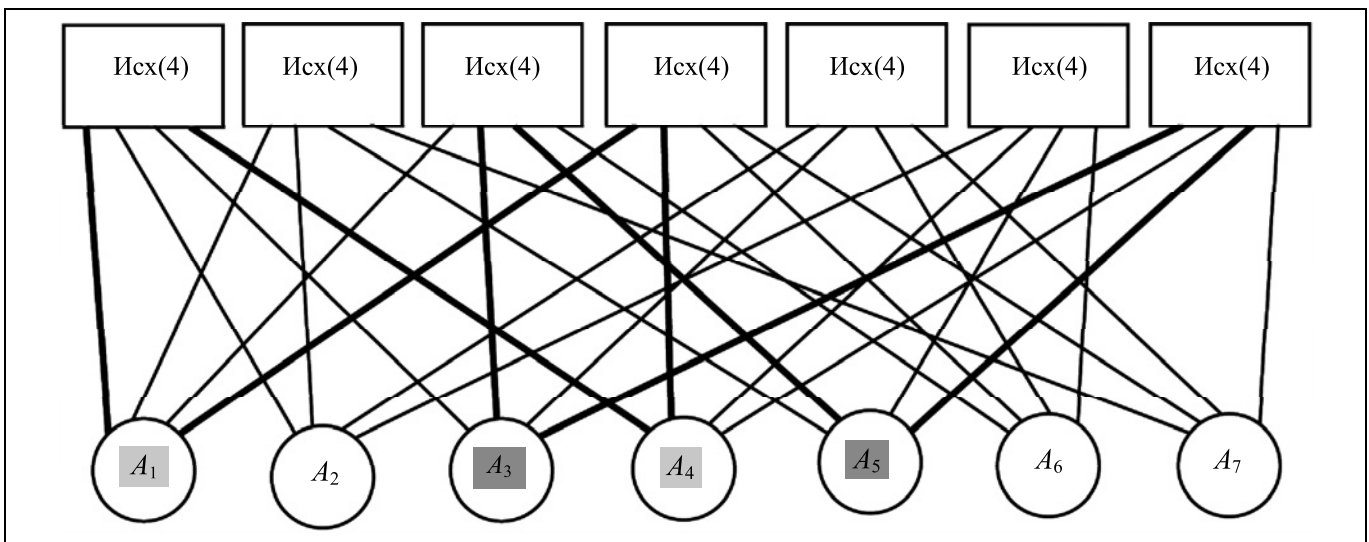


Рис. 3. Схема ПРС(7, 4, 2) в виде минимального квазиполного графа

ло дуплексных портов абонента до m . Берется N копий $\text{ИсхС}(K)$ с $K = rm$, и к ним подсоединяется $R = rN$ абонентов так, что каждый абонент соединяется с любым абонентом последовательно только через одну копию $\text{ИсхС}(K)$, и любая пара абонентов соединяется через σ разных сетей $\text{ИсхС}(K)$.

Таблица 3

Межсоединения в простейшей СС при $m = 4$ и $\sigma = 2$

Копии ИсхС(4)	Абоненты			
	1	2	3	4
1	1	2	3	4
2	1	2	5	7
3	1	3	5	6
4	1	4	6	7
5	2	3	6	7
6	2	4	5	6

Таблица межсоединений $\text{РасС}(R)$ формируется следующим образом. Различные копии $\text{ИсхС}(K)$ размещены по портам в N строках таблицы, а различные ПРС(N, m, σ) — в прямоугольных областях таблицы шириной в m столбцов. В табл. 4 приведен пример расширения сети $\text{ИсхС}(9)$ в сеть $\text{РасС}(21)$ на базе сети ПРС(7, 3, 1).

По построению в сети $\text{РасС}(R)$ любые два абонента, номера которых не совпадают по $\text{mod } N$, соединены друг с другом последовательно только через одну копию $\text{ИсхС}(K)$ и используют только ее маршрутные свойства. Наоборот, любые два абонента, номера которых совпадают по $\text{mod } N$, соединены друг с другом последовательно через m разных копий $\text{ИсхС}(K)$. Эти свойства обеспечивают сохранение в сети $\text{РасС}(R)$ маршрутных свойств сети $\text{ИсхС}(K)$. При этом образуются N подмно-

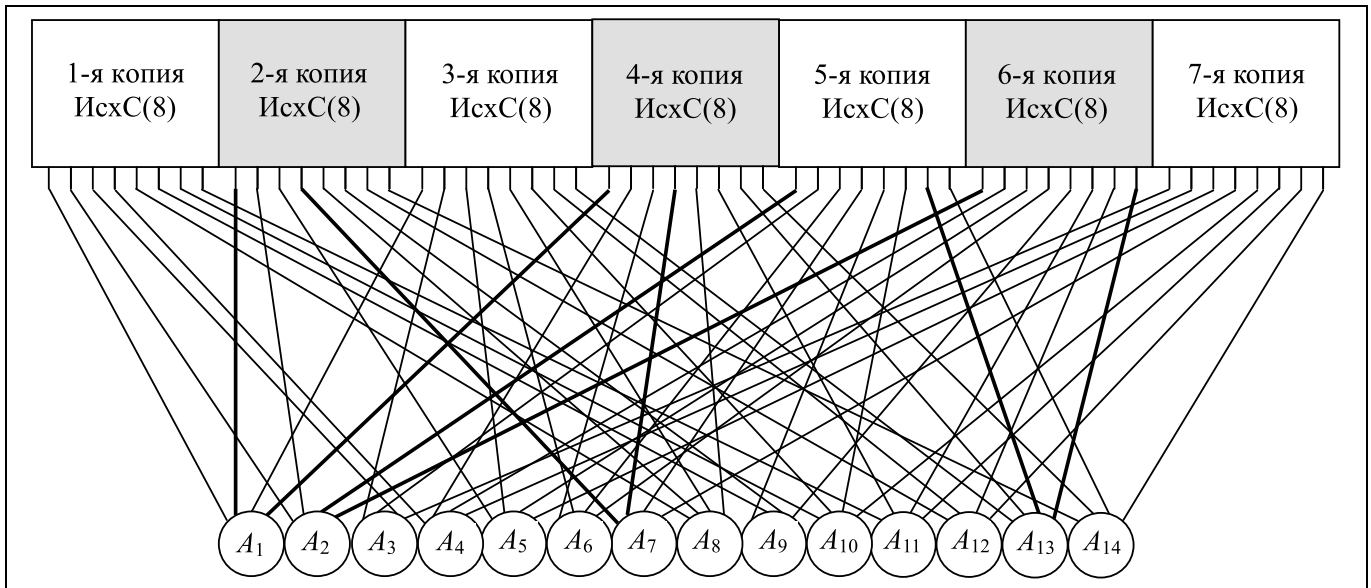


Рис. 4. Схема сети РасС(14) в виде двудольного графа

жеств по r абонентов, с увеличенной в m раз пропускной способностью в сети РасС(R) между ними.

Аналогично, в табл. 5 приведен пример расширения сети ИсхС(8) в сеть РасС(14) на базе сети ПРС(7, 4, 2). Схема сети Рас(14) представлена на рис. 4.

2. РАСШИРЕНИЕ СНИЗУ

Простейшая сеть «Ангара» (минимальная 1-мерная решетка) состоит из $M_1 = 2$ 24-портовых маршрутизаторов, объединяет $N_1 = 40$ процессоров, имеет диаметр $D_1 = 5$ скачков и сложность $S_1 = M_1$.

Расширим описанным выше методом 24-портовый маршрутизатор до 56-портового маршрутизатора с $n_1 = 56$ абонентами путем применения «хабов» 1×3 (рис. 5). Он имеет диаметр $d_1 = 3$ скачка и содержит $m_1 = 7$ 24-портовых маршрутизаторов и 56 «хабов». В дальнейшем рассматривается идеальный случай размещения «хабов» внутри абонентов. Поэтому принимаем, что расширенный маршрутизатор имеет сложность $s_1 = m_1$.

Для организации связи между такими маршрутизаторами в одномерном торе исключаются абоненты 1, 15, 29, 43 и 2, 16, 30, 44, а их 24 порта используются для связи слева и справа. При этом число подключаемых абонентов уменьшается на 8 (до 48). В двумерном торе дополнительно используются порты абонентов 8, 22, 36, 50 и 9, 23, 37, 51, а их 24 порта используются для связи сверху и снизу. Опять число подключаемых абонентов уменьшается еще на 8 (до 40).

Таблица 4

Межсоединения в сети РасС(21), состоящей из семи копий ИсхС(9) со встроенными в них тремя копиями ПРС(7, 3, 1)

Копии ИсхС(9)	Порты копий ИсхС(9)								
	1	2	3	4	5	6	7	8	9
	1-я ПРС			2-я ПРС			3-я ПРС		
1	1	7	5	8	14	12	15	21	19
2	2	1	6	9	8	13	16	15	20
3	3	2	7	10	9	14	17	16	21
4	4	3	1	11	10	8	18	17	15
5	5	4	2	12	11	9	19	18	16
6	6	5	3	13	12	10	20	19	17
7	7	6	4	14	13	11	21	20	18

Таблица 5

Межсоединения в сети РасС(14), состоящей из семи копий ИсхС(8) со встроенными в них двумя копиями ПРС(7, 4, 2)

Копии ИсхС(8)	Порты копий ИсхС(8)							
	1	2	3	4	5	6	7	8
	1-я ПРС (7, 4, 2)				2-я ПРС (7, 4, 2)			
1	1	2	3	4	8	9	10	11
2	1	2	5	7	8	9	12	14
3	1	3	5	6	8	10	12	13
4	1	4	6	7	8	11	13	14
5	2	3	6	7	9	10	13	14
6	2	4	5	6	9	11	12	13
7	3	4	5	7	10	11	12	14

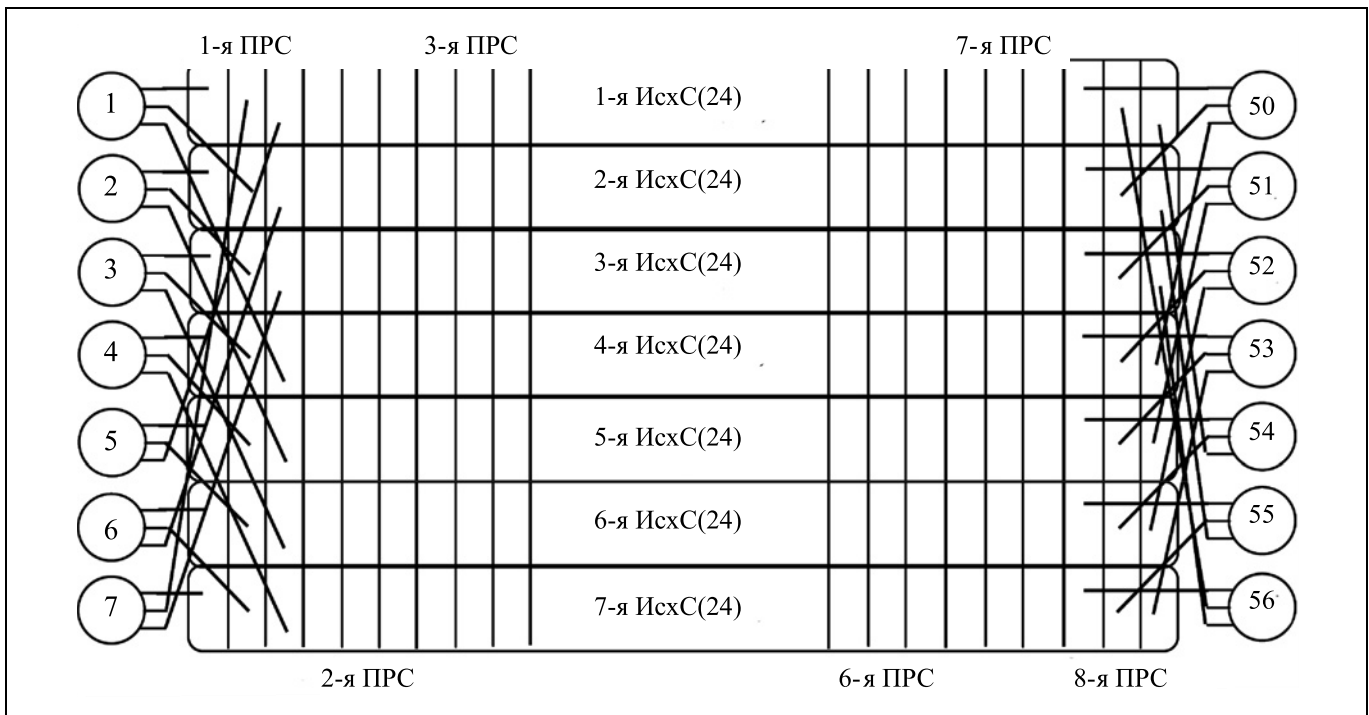


Рис. 5. Схема построения сети РасС(56) из семи копий ИсхС(24) со встроенными в них восемью копиями ПРС(7, 3, 1)

Составим кольцо (1-мерный тор) с «хабами» из m_2 56-портовых маршрутизаторов. Оно имеет диаметр в $d_2 = m_2/2 + 5$ скачков, объединяет $n_2 = 48m_2$ абонентов и имеет сложность $s_2 = 7m_2$.

Аналогичное кольцо сети «Ангара» состоит из M_2 24-портовых маршрутизаторов, имеет диаметр $D_2 = M_2/2 + 4$ скачков, объединяет $N_2 = 16M_2$ процессоров и имеет сложность $S_2 = M_2$.

При $m_2 = M_2$ получаем, что $n_2 = 3N_2$, $d_2 = D_2 + 1$ и $s_2 = 7S_2$, т. е. сеть с «хабами» содержит в 3 раза большее число абонентов, имеет почти равный диаметр и в семь раз большую сложность.

При равном числе процессоров $n_2 = N_2$ получаем, что $m_2 = M_2/3$, $d_2 \approx D_2/3 + 3,7$ и $s_2 = 7S_2/3$, т. е. сеть с «хабами» имеет несколько меньший диаметр, но и в 2,3 раза большую сложность. Размер диаметра более точно определяется при задании M_2 . При $M_2 = 8$ имеем $D_2 = 8$ и $d_2 \approx 6,3$. При $M_2 = 16$ имеем $D_2 = 12$ и $d_2 \approx 7,7$.

Наконец, при одинаковой сложности $s_2 = S_2$ получаем, что $m_2 = M_2/7$, $d_2 \approx D_2/7 + 4,24$ и $n_2 = 48M_2/7 = 3N_2/7$, т. е. сеть с «хабами» имеет несколько меньший диаметр, но и содержит в 2,3 раза меньшее число абонентов. При $M_2 = 8$ имеем $D_2 = 8$ и $d_2 \approx 5,7$. При $M_2 = 16$ имеем $D_2 = 12$ и $d_2 \approx 6,76$.

Составим теперь 2-мерный тор с «хабами» из m_3 56-портовых маршрутизаторов в каждом измерении. Он имеет диаметр в $d_3 = m_3 + 6$ скачков, объединяет $n_3 = 40m_3^2$ абонентов и имеет сложность $s_3 = 7m_3^2$.

Аналогичный 2-мерный тор сети «Ангара» из M_3 24-портовых маршрутизаторов в каждом измерении имеет диаметр $D_3 = M_3 + 5$ скачков, объединяет $N_3 = 8M_3^2$ процессоров и имеет сложность $S_3 = M_3^2$.

При $m_3 = M_3$ получаем, что $d_3 = D_3 + 1$, $n_3 = 5N_3$ и $s_3 = 7S_3$, т. е. сеть с «хабами» при почти равном диаметре содержит в 5 раз большее число абонентов, но и имеет в 7 раз большую сложность.

При равном числе процессоров $n_3 = N_3$ получаем, что $m_3 = M_3/\sqrt{5} = M_3/2,2$, $d_3 = D_3/2,2 + 4,7$ и $s_3 = 7m_3^2 = 7S_3/5$, т. е. сеть с «хабами» имеет несколько меньший диаметр, но и в 7/5 раз большую сложность. Размер диаметра более точно определяется при задании M_3 . Если $M_3 = 8$ и $D_3 = 13$, то $d_3 \approx 8,1$. Если же $M_3 = 16$ и $D_3 = 21$, то $d_3 \approx 11,7$.

Наконец, при одинаковой сложности $s_3 = S_3$ получаем, что $m_3 = M_3/\sqrt{7}$, $d_3 = D_3/2,6 + 4,9$ и



$n_3 = 40(M_3/\sqrt{7})^2 = 5N_3/7$, т. е. сеть с «хабами» имеет несколько меньший диаметр и содержит в 7/5 раз меньшее число процессоров. Если $M_3 = 8$ и $D_3 = 13$, то $d_3 \approx 9,9$. Если же $M_3 = 16$ и $D_3 = 21$, то $d_3 \approx 13$.

В табл. 6 сравниваются характеристики сети «Ангара» с «хабами» 1×3 и без них. Расширенный маршрутизатор сравнивается с минимальной 1-мерной решеткой. Видно, что диаметр сети можно уменьшить, а число абонентов увеличить в несколько раз путем некоторого увеличения сложности сети.

Дальнейшее улучшение характеристик сети можно получить благодаря применению расширенного маршрутизатора в 3- и 4-мерных торах. Сеть «Ангара» не может иметь топологию 3-мерного тора, так как в ней не хватает портов.

Для организации связи между расширенными маршрутизаторами в 3-мерном торе исключаются еще 8 абонентов, а их 24 порта используются для связи вперед и назад. При этом число подключаемых абонентов уменьшается еще на 8 (до 32).

Составим теперь 3-мерный тор из m_4 расширенных маршрутизаторов в каждом измерении. Он имеет диаметр $d_4 = 3m_4/2 + 7$ скачков и объединяет

$$n_4 = 32m_4^3 \text{ абонентов.}$$

Таблица 6

Сравнительные характеристики сети «Ангара» с внутренними «хабами» при расширении снизу

Исходная сеть «Ангара»	D	N	S	$s = S/N$
Расширенный маршрутизатор	$D/2$	$7N/3$	$7S/2$	$3s/2$
1-мерный тор с расширенным маршрутизатором	$M_2 = m_2$ $D + 1$ $\sim D/3 + 3,7$ $\sim D/7 + 4,2$	$3N$ N $3N/7$	$7S$ $7S/3$ S	$7s/3$ $7s/3$ $7s/3$
2-мерный тор с расширенным маршрутизатором	$M_3 = m_3$ $D + 1$ $\sim D/2,2 + 4,7$ $\sim D/2,6 + 4,9$	$5N$ N $5N/7$	$7S$ $7S/5$ S	$7s/5$ $7s/5$ $7s/5$

Таблица 7

Характеристики сети «Ангара» в 3- и 4-мерном торе

3-мерный тор			4-мерный тор		
m_4	N_4	D_4	m_5	N_5	D_5
3	864	10	3	1944	12
4	2048	13	4	6144	16
5	4000	14,5	5	15 000	18
6	6912	16	6	31 104	20

Аналогично составим 4-мерный тор из m_5 расширенных маршрутизаторов в каждом измерении. Он имеет диаметр $d_5 = 2m_5 + 8$ скачков и объединяет $n_5 = 24m_5^4$ абонентов.

Характеристики этих торов приведены в табл. 7.

В рамках данного подхода можно создать расширенный коммутатор и в отдельном корпусе на 36, 56 и 78 портов. Для этого достаточно взять 3, 7, и 13 сцепок 8-портовых маршрутизаторов совместно с 24 «хабами» 1×2 , 1×3 и 1×4 , размещенными между сцепками маршрутизаторов и внешними портами корпуса. Такой расширенный «хаб» имеет большую сложность и энергопотребление, но позволяет применять процессоры без «хабов». Такие сложные расширенные маршрутизаторы дают возможность дополнительно сокращать диаметры сетей или увеличивать число процессоров, когда применяются абоненты с «хабами».

3. РАСШИРЕНИЕ СВЕРХУ

Расширим минимальную решетку «Ангара», применяя «хабы» 1×3 . Эта решетка состоит из $M_1 = 2$ 24-портовых маршрутизаторов, имеет диаметр $D_1 = 6$ скачков, одержит $N_1 = 40$ процессоров и имеет сложность $S_1 = M_1$. В результате расширения получим сеть с «хабами», которая состоит из $m_1 = 14$ 24-портовых коммутаторов, имеет диаметр $d_1 = 7$ скачков, содержит $N_1 = 91$ абонентов и имеет сложность $s_1 = 7S_1$.

Аналогично расширим 1-мерный тор «Ангара», который состоит из M_2 24-портовых маршрутизаторов, имеет диаметр в $D_2 = M_2/2 + 4$ скачков, объединяет $N_2 = 16M_2$ процессоров и имеет сложность $S_2 = M_2$.

Расширенная сеть с «хабами» состоит из 7 торов с m_2 24-портовыми маршрутизаторами каждый, имеет диаметр $d_2 = m_2/2 + 5$ и содержит $n_2 = 7 \lfloor 16m_2/3 \rfloor$ процессоров и имеет сложность $s_2 = 7m_2$.

Если $m_2 = M_2$, то $d_2 = D_2 + 1$, $n_2 = 7 \lfloor N_2/3 \rfloor \approx 7N_2/3$ и $s_2 = 7S_2$.

Если $n_2 = N_2$, то $m_2 \approx 3M_2/7$, $d_2 \approx 3D_2/7 + 3,3$ и $s_2 = 3S_2$. При $M_2 = 8$ имеем $D_2 = 8$ и $d_2 \approx 6,7$, а при $M_2 = 16$ имеем $D_2 = 12$ и $d_2 \approx 8,3$.

Если $s_2 = S_2$, то $m_2 = M_2/7$, $d_2 \approx D_2/7 + 4,4$ и $n_2 = 7 \lfloor 16m_2/3 \rfloor \approx N_2/3$. При $M_2 = 8$ имеем $D_2 = 8$ и $d_2 \approx 5,5$, а при $M_2 = 16$ имеем $D_2 = 12$ и $d_2 \approx 6,1$.

Аналогично расширим 2-мерный тор «Ангара». Расширенная сеть с «хабами» состоит из семи 2-мерных торов, каждый из которых имеет m_3

24-портовых маршрутизаторов в каждом измерении. Расширенная сеть имеет диаметр в $d_3 = m_3 + 6$ скачков, объединяет $n_3 = 7\lfloor 8m_3^2/3 \rfloor$ абонентов и имеет сложность $s_3 = 7m_3^2$.

Исходный 2-мерный тор сети «Ангара» состоит из M_3 24-портовых маршрутизаторов в каждом измерении, имеет диаметр в $D_3 = M_3 + 5$ скачков, объединяет $N_3 = 8M_3^2$ процессоров и имеет сложность $S_3 = M_3^2$.

Если $m_3 = M_3$, то $d_3 = D_3 + 1$, $n_3 = 7\lfloor N_3/3 \rfloor \approx 7N_3/3$ и $s_3 = 7S_3$.

Если $n_3 = N_3$, то $m_3 \approx \sqrt{(3/7)M_3} = 0,65M_3$, $d_3 \approx 0,65D_3 + 2,75 = D_3/1,5 + 2,75$ и $s_3 = 3S_3$. При $M_3 = 8$ имеем $D_3 = 13$ и $d_3 \approx 11,2$, а при $M_3 = 16$ имеем $D_3 = 21$ и $d_3 \approx 16,4$.

Если $s_3 = S_3$, то $m_3 = M_3/\sqrt{7}$, $d_3 \approx D_3/2,6 + 4,1$ и $n_3 = 7\lfloor 8m_3^2/3 \rfloor \approx N_3/3$. При $M_3 = 8$ имеем $D_3 = 13$ и $d_2 \approx 9,1$, а при $M_2 = 16$ имеем $D_2 = 21$ и $d_2 \approx 12,2$.

Таблица 8

Сравнительные характеристики сети «Ангара» с внутренними «хабами» 1×3 при расширении сверху

Исходная сеть «Ангара»	D	N	S	$s = S/N$
Расширение минимальной сети	$7D/6$	$9N/4$	$7S$	$\sim 3s$
Расширенный 1-мерный тор	$M_2 = m_2$ $D + 1$	$7N/4$	$7S$	$4s$
	$\sim 3D/7 + 3,3$	N	$3S$	$3s$
	$\sim D/7 + 4,4$	$N/3$	S	$3s$
Расширенный 2-мерный тор	$M_3 = m_3$ $D + 1$	$7N/3$	$7S$	$3s$
	$\sim D/1,5 + 2,4$	N	$3S$	$3s$
	$\sim D/2,6 + 3,5$	$N/3$	S	$3s$

Таблица 9

Диаметры сети «Ангара» с внутренними «хабами» 1×3 при расширении снизу и сверху

D	1-мерный тор		2-мерный тор	
	8	12	13	21
Снизу ($n = N$)	6,3	7,7	8,1	11,7
Снизу ($s = S$)	5,7	6,8	9,9	13
Сверху ($n = N$)	6,7	8,3	11,2	16,4
Сверху ($s = S$)	5,5	6,1	9,1	12,2

Таблица 10

Сравнительные характеристики сети 1-отказоустойчивой сети «Ангара» с внутренними «хабами» 1×4

Исходная сеть «Ангара»	D	N	S	$s = S/N$
Расширение минимальной сети	$7D/6$	$7N/4$	$7S$	$4s$
Расширенный 1-мерный тор	$M_2 = m_2$ $D + 1$	$7N/4$	$7S$	$4s$
	$4D/7 + 3,7$	N	$4S$	$4s$
	$D/7 + 4,4$	$N/4$	S	$4s$
Расширенный 2-мерный тор	$M_3 = m_3$ $D + 1$	$7N/4$	$7S$	$4s$
	$\sim D/1,33 + 2,25$	N	$4S$	$4s$
	$\sim D/2,6 + 4,1$	$N/4$	S	$4s$

В табл. 8 сравниваются характеристики сети «Ангара» с «хабами» 1×3 и без них при расширении сверху.

В табл. 9 сравниваются диаметры d торов при расширении снизу и сверху. Табл. 9 не выявляет превосходства по d какого-либо подхода к расширению сети. Однако сравнение табл. 7 и 8 показывает меньшую в 1,7 раза удельную сложность при расширении снизу.

Однако при расширении сверху можно создавать канално-отказоустойчивые системные сети. В частности, если взять сеть ПРС(7, 4, 2), задаваемую табл. 3 и рис. 3, то можно создать 1-отказоустойчивую сеть с «хабами», представленную в табл. 10.

Расширим минимальную решетку «Ангара», применяя «хабы» 1×4. В результате расширения получим сеть с «хабами», которая имеет диаметр $d_1 = 7$ скачков, содержит $n_1 = 70$ абонентов, состоит из $m_1 = 14$ 24-портовых коммутаторов и является 1-отказоустойчивой по каналам.

Аналогично расширим 1-мерный тор «Ангара», применяя «хабы» 1×4. Расширенная 1-отказоустойчивая сеть с «хабами» состоит из 7 торов с m_2 24-портовыми маршрутизаторами, имеет диаметр $d_2 = m_2 + 5$, содержит $n_2 = 7(16m_2/4) = 28m_2$ абонентов и является 1-отказоустойчивой по каналам. Она имеет сложность $s_2 = 7m_2$.

Если $m_2 = M_2$, то $d_2 = D_2 + 1$ и $n_2 = 7N_2/4$. При этом сложность отказоустойчивой сети с «хабами» составляет $s_2 = 7S_2$.

Если $n_2 = N_2$, то $m_2 = 4M_2/7$, $d_2 = 4D_2/7 + 3,7$ скачков и $s_2 = 4S_2$. При $M_2 = 8$ имеем $D_2 = 8$ и $d_2 \approx 8,3$, а при $M_2 = 16$ имеем $D_2 = 12$ и $d_2 \approx 10,6$.



Таблица 12

Характеристики гипотетической гиперкубовой сети в формате R/D

Число маршрутизаторов p в каждом измерении	4	8
2-мерный гиперкуб	576/6	1792/6
3-мерный гиперкуб	2112/9	10752/9
4-мерный гиперкуб	7680/12	57344/12

Если $s_2 = S_2$, то $m_2 = M_2/7$, $d_2 = D_2/7 + 4,4$ и $n_2 = 7N_2/4$. При $M_2 = 8$ имеем $D_2 = 8$ и $d_2 \approx 5,5$, а при $M_2 = 16$ имеем $D_2 = 12$ и $d_2 \approx 6,1$.

Аналогично расширим 2-мерный тор «Ангарты». Расширенная 1-отказустойчивая сеть с «хабами» состоит из семи 2-мерных торов, каждый из которых имеет m_3 24-портовых маршрутизаторов в каждом измерении. Расширенная 1-отказустойчивая сеть имеет диаметр в $d_3 = m_3 + 6$ скачков, объединяет $n_3 = 7(16m_3^2)/4$ абонентов и имеет сложность $s_3 = 7m_3^2$.

Если $m_3 = M_3$, то $d_3 = D_3 + 1$, $n_3 = 7N_3/4$ и $s_3 = 7S_3$.

Если $n_3 = N_3$, то $m_3 \approx \sqrt{(4/7)} M_3 = 0,75M_3$, $d_3 = 0,75D_3 + 2,25$ и $s_3 = 3S_3$. При $M_3 = 8$ имеем $D_3 = 13$ и $d_3 \approx 12$, а при $M_3 = 16$ имеем $D_3 = 21$ и $d_3 \approx 18$.

Если $s_3 = S_3$, то $m_3 = M_3/\sqrt{7} \approx M_3/2,6$, $d_3 = D_3/2,6 + 4,1$ и $n_3 = 7(8m_3^2)/4 = N_3/4$. При $M_2 = 8$ имеем $D_2 = 13$ и $d_2 \approx 9,1$, а при $M_2 = 16$ имеем $D_2 = 21$ и $d_2 \approx 12,2$.

4. СИСТЕМНЫЕ СЕТИ СО СЛОЖНЫМ РАСШИРЕННЫМ МАРШРУТИЗАТОРОМ

Составим 18-портовый маршрутизатор из сцепки трех 8-портовых маршрутизаторов. Он имеет внутренний диаметр в 3 скачка и проходную задержку в 1 скачок при трех каналах для соединения таких маршрутизаторов.

Расширим этот маршрутизатор в однокорпусной сложный 42-портовый маршрутизатор с внутренними «хабами». Он содержит 7 сцепок 8-портовых маршрутизаторов и 42 «хаба» 1×3 . Он имеет в 5–6 раз большие сложность и энергопотребление, чем 24-портовый маршрутизатор, но и позволяет строить сети с более высокими характеристиками по масштабированию и быстродействию.

В табл. 11 приводятся характеристики сетей из 42-портовых маршрутизаторов с торовой топологией.

Таблица 11

Характеристики гипотетической торовой сети в формате R/D (число узлов/диаметр)

Число маршрутизаторов в измерении	4	8	12	16
1-мерный тор	144/5	288/7	288/9	576/11
2-мерный тор	576/8	1920/12	4320/16	7680/20
3-мерный тор	1536/11	12048/14	41472/17	98304/23

При применении 42-портового маршрутизатора для построения p -ичного гиперкуба его внутренний диаметр и проходная задержка составляет 3 скачка. В табл. 12 приведены характеристики таких гиперкубов.

Если же расширять 24-портовый маршрутизатор, используя «хабы» 1×4 , то расширенный сложный маршрутизатор будет содержать те же 42 порта, но обретет канальную отказоустойчивость. На его основе можно строить 1-отказоустойчивые p -ичные гиперкубы с числом процессоров из табл. 11, но с диаметром в $4/3$ раза большим.

ЗАКЛЮЧЕНИЕ

Предложен способ расширения возможностей сети «Ангара» путем совместного применения собственных 24-портовых маршрутизаторов и рыночных «хабов» 1×3 и 1×4 . В различных вариантах их применения имеется возможность увеличения масштабируемости сети (повышения числа процессоров в несколько раз), быстродействия сети (сокращения диаметра примерно в два раза) и ее канальной отказоустойчивости. Кроме того, появляется возможность создания сложных расширенных сложных маршрутизаторов с еще большим числом портов (36, 42, 56, 78) в одном корпусе, что, в свою очередь, позволяет создавать системные сети в виде многомерных торов и обобщенных гиперкубов с еще более высокими характеристиками. При этом можно строить 1-отказоустойчивые обобщенные гиперкубы из 1-отказоустойчивых расширенных сложных маршрутизаторов с 42 портами на базе «хабов» 1×4 .

ЛИТЕРАТУРА

1. Alverson, R., Roweth, D. and Kaplan, L. The Gemini System Interconnect // 18th IEEE Symposium on High Performance Interconnects. — 2009. — P. 3–87.
2. Alverson, R., Froese, E., Kaplan, L. and Roweth, D. Cray XC® Series Network // URL: <https://www.cray.com/sites/default/files/resources/CrayXCNetwork.pdf>.
3. Михеев В.А., Симонов А.С., Слуцкий А.И. и др. Реализация высокоскоростной сети для суперкомпьютерных систем: проблемы, результаты, развитие // URL: <http://2013.nscf.ru/>

- TesisAll/Section%201/12_2761_SiNonovAS_S1.pdf. [Mikheev, V.A., Simonov, A.S., Slutskii, A.I., et al. Realizatsiya vysokoskorostnoi seti dlya superkomp'yuternykh sistem: problemy, rezultaty, razvitiye. (In Russian)]
4. Подлазов В.С. Возможности кольцевых каналов в масштабируемых многопроцессорных вычислительных системах с общей разделяемой памятью // Тр. Института проблем управления РАН. — 1999. — Т. VI. — М.: ИПУ РАН. — С. 91—99. [Podlazov, V.S. *Vozmozhnosti kol'tsevykh kanalov v masshtabiruemykh mnogoprotseornykh vychislitel'nykh sistemakh s obshchei razdelyaemoi pamyat'yu* // Trudy instituta problem upravleniya RAN. — 1999. — Vol. VI. — М.: IPU RAN. — P. 91—99. (In Russian)]
 5. Каравай М.Ф., Подлазов В.С. Системная сеть с малым диаметром из малопортовых маршрутизаторов // Управление большими системами. — 2015. — Вып. 56. — М.: ИПУ РАН. — С. 201—210. [Karavay, M.F., Podlazov, V.S. *Small-Diameter System Area Network Composed of Small-Port Routers* // Automation and Remote Control. — 2017. — Vol. 78, no. 9. — P. 1674—1679.]
 6. Симонов А.С., Макагон Д.В., Жабин И.А. и др. Первое поколение высокоскоростной коммуникационной сети «Ангара» // Научные технологии. — 2014. — Т. 15, № 1. — С. 21—28. [Simonov, A.S., Makagon, D.V., Zhabin, I.A., et al. *Pervoe pokolenie vysokoskorostnoi kommunikatsionnoi seti «Angara»* // Naukoemkie tekhnologii. — 2014. — Vol. 15, no. 1. — P. 21—28. (In Russian)]
 7. Stegailov, V., Agarkov, A., Biryukov, S., 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. — Springer — Cham. — 2017. — P. 327—336.
 8. Агарков А.А., Исмаилов Т.Ф., Макагон Д.В. и др. Предварительные результаты оценочного тестирования отечественной высокоскоростной коммуникационной сети «Ангара» // Параллельные вычислительные технологии (ПАВТ'2016): тр. междунар. науч. конф., Архангельск, 2016. — Челябинск: Изд. центр ЮУрГУ, 2016. — С. 42—53. [Agarkov, A.A. *Ismailov, T.F., Makagon, D.V., et al. Predvari-*
- tel'nye rezultaty otsenochnogo testirovaniya otechestvennoi vysokoskorostnoi kommunikatsionnoi seti «Angara» // Parallel'nye vychislitel'nye tekhnologii (PAVT'2016): tr. mezhdunar. nauch. konf., Arkhangel'sk, 2016. — Chelyabinsk: Izd. tsentr YuUrGU, 2016. — S. 42—53. (In Russian)]
9. 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.C. *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.]
 10. Холл М. Комбинаторика. — М.: Мир, 1970. — 424 с. [Hall, M. *Combinatorial Theory*. — Waltham, Mass: Blaisdell Publishing Company, 1967. — 320 p.]
 11. 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.C. *Combinatorial methods for constructing bipartite uniform minimal quasicomplete graphs (symmetrical block designs)* // Automation and Remote Control. — 2009. — Vol. 70, no. 2. — P. 312—327.]

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

Поступила в редакцию 4.02.2019, после доработки 13.05.2019.
Принята к публикации 2.12.2019.

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

Каравай Михаил Федорович — д-р техн. наук,
✉ mkaravay@ipu.ru,

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

EXPANDING THE CAPABILITIES OF THE ANGARA SYSTEM AREA NETWORK

V.S. Podlazov, M.F. Karavay

V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

✉ podlazov@ipu.ru

Abstract: A method for expanding the capabilities of the Angara system area network through the sharing of its own 24-port routers and market hubs 1×3 and 1×4 is considered. The method is based on the use of networks with the topology of quasi-complete graphs constructed in an extended elemental basis consisting of switches (routers) and hubs of duplex channels. These networks are the basis of the method of invariant expansion of arbitrary networks, which preserves transmission delays and increases the number of subscribers. Using this method, a distributed extended router is constructed from Angara network routers, and networks based on it are constructed with the structure of 1, 2, 3, 4-dimensional tori, either containing a significantly larger number of subscribers with the same network diameter, or having a smaller diameter with the same number of subscribers compared to the Angara network. In fact, the method under consideration made it possible to increase the number of processors, reduce the diameter and increase the channel fault tolerance of the network.

Keywords: system networks of supercomputers, Angara network, networks with topology of quasi-complete graphs, invariant expansion of networks, number of network subscribers and network diameter, scaling and performance.