

# РЕГУЛЯРНЫЕ ПЛОСКИЕ СЕТИ ДЛЯ СУПЕРКОМПЬЮТЕРОВ

В. С. Подлазов, Г. Г. Стецюра

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

Предложен способ построения плоских коммутируемых локальных сетей для создания дешевых суперкомпьютеров из персональных компьютеров.

## ВВЕДЕНИЕ

В последнее десятилетие в университетской среде получила тенденция построения суперкомпьютеров с использованием коммутируемой сети (Fast)Ethernet на базе системного программного обеспечения кластера Beowulf [1–3]. Для обеспечения высокой пропускной способности и сокращения задержек было предложено строить эту сеть как плоскую сеть (ПС), в которой путь между любым двумя персональными компьютерами (станциями) проходит только через один сетевой коммутатор [2, 3]. Для разных станций это могут быть разные коммутаторы. В каждой станции может быть 2–6 (обычно 3–4) сетевых карты, которые подключаются к разным коммутаторам.

Для выбора размера и числа коммутаторов, числа сетевых карт и для построения схемы подсоединения станций к коммутаторам в работах [2, 3] применялись генетические алгоритмы. Опыт автора по применению генетических алгоритмов для оптимизации структуры мультикольцевых сетей [4, 5] показал, что они часто не находят решений с сильно выделяющимися значениями целевой функции. Обычно уникальные или регулярные решения характеризуются такой целевой функцией. Исходя из этого опыта, в статье предлагается регулярный способ построения плоских и разреженных плоских сетей [6] с заранее предсказуемыми характеристиками, которые оказываются лучше, чем в кластерах KLAT2 и KASY0, построенных по генетическим алгоритмам. В разреженной плоской сети (РПС) путь между любыми двумя станциями проходит не более чем через один сетевой коммутатор 2-го уровня, и для значительной доли станций он проходит только через один коммутатор 1-го уровня.

Для анализа свойств ПС и РПС составлена простая программа, которая по схеме подсоединения станций к коммутаторам в некоторой сети выдает различные ее па-

раметры и характеристики. Все используемые в данной работе свойства сетей получены с помощью данной программы, и они даются без теоретических доказательств.

## 1. ЗАМКНУТЫЕ ПЛОСКИЕ СЕТИ

Рассмотрим ПС с точки зрения теории коммутации, считая сетевой коммутатор самомаршрутизируемым неблокируемым коммутатором. Введем следующие обозначения:  $N$  — число станций,  $m$  — число сетевых карт в каждой станции,  $M$  — число сетевых коммутаторов,  $K$  — число портов в каждом коммутаторе.

В ПС полный коммутатор  $N \times N$  заменяется на однокаскадный набор из  $M$  полных коммутаторов  $K \times K$  ( $K < N$ ), который остается самомаршрутизируемым и неблокируемым. Это достигается путем увеличения числа выходов каждой станции с 1 до  $m$  и необходимого распределения этих выходов по  $M$  коммутаторам.

Простейший регулярный способ построения ПС состоит в следующем. Все  $N$  станций разделяются на  $L$  равных групп —  $G_1, \dots, G_L$  по  $n$  станций в группе. Все станции любой пары групп  $G_i$  и  $G_j$  ( $i \neq j$ ) подсоединяются к одному коммутатору, поэтому  $n = K/2$ . Все станции одной группы подсоединяются к  $m$  разным коммутаторам. Для обеспечения неблокируемости достаточно, чтобы число групп и число коммутаторов задавались выражениями  $L = m + 1$  и  $M = L(L - 1)/2$ . При этом должно выполняться соотношение  $Nm \leq KM$ . Если в последнем выражении имеет место равенство, то для распространенных коммутаторов возможны ПС с параметрами и характеристиками, приведенными в табл. 1.

В работах [2, 3] для расчета удельной полосы пропускания, приходящейся на пару соединенных станций, используется следующая формула:

$$w = \nu M(K - 1)K/N(N - 1) = \nu m(K - 1)/(N - 1),$$



где  $v = 200$  Мбит/с — полоса одной сетевой карты при дуплексном режиме работы в сети FastEthernet. Общая полоса пропускания  $W$  плоской сети в  $N$  раз больше, т. е.

$$W = vMK(K - 1)/(N - 1). \quad (1)$$

Правильность формулы (1) подтверждена в работах [2, 3] измерениями на тестовых задачах.

Однако не следует забывать, что рассмотренное разбиение коммутатора  $N \times N$  на набор из  $M$  коммутаторов  $K \times K$  увеличивает общую сложность сети. Пусть сложность коммутатора измеряется в числе точек коммутации. Сложность коммутатора  $N \times N$  составляет величину  $S_1 = N^2$ . Суммарная сложность  $M$  коммутаторов  $K \times K$   $S_M = MK^2 = 2N^2(L - 1)/L = 2S_1(1 - 1/L)$ . Кроме того, добавляется  $N$  выходных мультиплексоров сложности  $m$ , сложность которых  $Nm = N(L - 1) \ll S_M$ . Замечательно, что сложность увеличивается не более, чем вдвое.

Однако, построенная таким образом ПС обладает рядом недостатков, которые делают ее непригодной для построения кластеров Beowulf на сети (Fast)Ethernet.

Прежде всего, в ней задействованы все порты всех коммутаторов, что делает ее замкнутой. При этом становится невозможным подключение специализированных станций (станций Internet, станций памяти и т. п.).

Таблица 1

Варианты замкнутых плоских сетей

$m$	$K$	$N$	$M$	$W$ , Гбит/с
2	16	24	3	6,3
	24	36		9,5
	32	48		12,7
	48	72		19,1
	96	144		38,3
3	16	32	6	9,3
	24	48		14,1
	32	64		18,9
	48	96		28,5
	64	128		38,1
4	16	40	10	12,3
	24	60		18,7
	32	80		25,1
	48	120		37,9
	64	160		50,7
5	16	48	15	15,3
	24	72		23,3
	32	96		31,3
	48	144		47,3
	64	192		63,3
6	16	56	21	18,3
	24	84		27,9
	32	112		37,5
	48	168		56,7
	64	224		75,9
	96	336		114,3

Далее, сеть Ethernet широко использует широковещание на канальном уровне. В коммутируемой сети для его реализации необходимо дерево коммутаторов, для создания которого требуются дополнительные порты. В замкнутой сети их нет.

И наконец, табл. 1 показывает ограниченную наращиваемость замкнутых ПС при  $N > 256$ . Особенно это справедливо в дешевых вариантах с  $m = 3-4$  и  $K = 24-32$ . Значительно лучшей наращиваемостью обладают разреженные плоские сети [6], для построения которых рассмотренный способ не подходит.

Требуется разработка другого регулярного способа построения ПС, который применим для построения и РПС. Этот способ рассматривается в следующих разделах.

## 2. ОТКРЫТЫЕ ПЛОСКИЕ СЕТИ

Способ построения ПС основан на разработке открытой сети, которую можно наращивать до нужного числа станций путем увеличения как числа коммутаторов, так и числа портов в них.

Открытая сеть 1-го ранга состоит из  $M_{1,m}$  коммутаторов с  $K_{1,m} = m$  портами. В ней каждая станция должна иметь пути, проходящие через один коммутатор, к максимально возможному числу  $m(m - 1)$  других станций. Открытая сеть с минимально возможным числом станций  $N_{1,m} = M_{1,m}$  называется простейшей сетью.

Рассмотрим подробнее свойства простейшей сети 1-го ранга. Будем называть порожденной группой  $PG(J)$  множество станций, с которыми станция  $J$  имеет соединение через один коммутатор. Будем включать в эту группу и саму станцию  $J$ . Обозначим через  $n_{1,m}(J)$  число станций в группе  $PG(J)$ , а через  $L_{1,m}$  — число разных по составу станций групп. Если в простейшей сети 1-го ранга  $n_{1,m}(J) = n_{1,m} = N_{1,m}$  и  $L_{1,m} = 1$ , то простейшая сеть 1-го ранга является плоской, и в ней  $N_{1,m} = m(m - 1) + 1$ .

Для  $m = 3$  удалось построить простейшую сеть 1-го ранга для  $N_{1,3} = M_{1,3} = 7$  с  $n_{1,3} = 7$  и  $L_{1,3} = 1$ . Конфигурация такой сети задается в табл. 2, где  $SW$  — номера коммутаторов,  $P$  — номера портов. Первый столбец задает номера коммутаторов, а первая строка — номера портов в них. В клетках таблицы задаются номера станций. Номера станций в третьем и четвертом столбцах получены циклическим сдвигом номеров станций в столбце 2 на 1 и 3 строки вниз.

Для  $m = 4$  удалось построить простейшую сеть 1-го ранга для  $N_{1,4} = M_{1,4} = 13$  с  $n_{1,4} = 13$ . Конфигурация такой сети задается в табл. 3. Номера станций в третьем и четвертом столбцах получены циклическим сдвигом номеров станций во втором столбце на 1, 3 и 9 строк вниз, соответственно.

Простейшие сети, представленные в табл. 2 и 3, являются плоскими.

Из открытой сети 1-го ранга можно построить открытую сеть 2-го ранга, увеличив число портов в коммутаторах и число станций до величин  $K_{2,m} = 2m$  и  $N_{2,m} = 2N_{1,m}$  соответственно. При этом число коммутаторов остается прежним:  $M_{2,m} = M_{1,m}$ . Обозначим через  $SW_i(J)$  и  $P_i(J)$  коммутатор и порт в нем, к которому подсоединяется станция  $J$  ( $N_{1,m} \leq J < N_{2,m}$ ) через  $i$ -ю карту. В открытой сети 2-го ранга подсоединение стан-

Таблица 2

**Простейшая сеть 1-го ранга при  $m = 3$** 

SW	P		
	0	1	2
0	0	6	4
1	1	0	5
2	2	1	6
3	3	2	0
4	4	3	1
5	5	4	2
6	6	5	3

Таблица 3

**Простейшая сеть 1-го ранга при  $m = 4$** 

SW	P			
	0	1	2	3
0	0	12	10	4
1	1	0	11	5
2	2	1	12	6
3	3	2	0	7
4	4	3	1	8
5	5	4	2	9
6	6	5	3	10
7	7	6	4	11
8	8	7	5	12
9	9	8	6	0
10	10	9	7	1
11	11	10	8	2
12	12	11	9	3

Таблица 4

**Простейшая сеть 2-го ранга при  $m = 3$** 

SW	P					
	0	1	2	3	4	5
0	0	6	4	7	13	11
1	1	0	5	8	7	12
2	2	1	6	9	8	13
3	3	2	0	10	9	7
4	4	3	1	11	10	8
5	5	4	2	12	11	9
6	6	5	3	13	12	10

Таблица 5

**Новые варианты плоских сетей**

$m$	$K$	$N$	$M$
3	24	56	7
	33	77	
	48	112	
	63	147	
	96	224	
4	24	78	13
	32	104	
	48	156	
	64	208	
	96	312	

ции  $J$  выполняется с соблюдением условий  $SW_i(J) = SW_i(J - N_{1,m})$  и  $P_i(J) = m + P_i(J - N_{1,m})$ .

В табл. 4 показана открытая сеть 2-го ранга при  $m = 3$ .

Эта сеть является простейшей сетью 2-го ранга, так как она получена расширением простейшей сети 1-го ранга. Для нее выполняется соотношение  $n_{2,3} = n_{1,3} + N_{1,3} = N_{2,3} = 14$  и  $L_{2,3} = 1$ . Поэтому она также является плоской.

Аналогично строится открытая сеть  $r$ -го ранга из сети  $(r - 1)$ -го ранга. В ней  $N_{r,m} = K_{r,m} = rm$  и  $M_{r,m} = M_{1,m}$ . В открытой сети  $r$ -го ранга подключение станции  $J$  ( $N_{r-1,m} \leq J < N_{r,m}$ ) выполняется с соблюдением условий  $SW_i(J) = SW_i(J - N_{r-1,m})$  и  $P_i(J) = m + P_i(J - N_{r-1,m})$ . При этом простейшая сеть переходит в простейшую, для которой выполняется соотношение  $n_{r,m} = n_{r-1,m} + N_{r-1} = N_r$ . Поэтому все такие сети плоские. В табл. 5 показаны плоские замкнутые сети, полученные этим способом для  $m = 3$  и  $m = 4$ . В результате мы получили ПС с меньшим числом портов в коммутаторах и (или) большим числом станций, чем указанные в табл. 1, при большем числе коммутаторов.

В большинстве ПС (см. табл. 5) используются все порты коммутаторов и не допускается подключение коммутатора 2-го уровня. Однако один или два порта в каждом коммутаторе легко освободить, уменьшив число станций. В табл. 6 эта процедура показана на примере ПС с  $K = 24$  порта, жирным шрифтом выделены удаляемые станции. Оставшиеся станции перенумеровываются и подключаются к портам 21 и 22. В результате удаляется три станции и остается по одному незанятому порту в трёх коммутаторах.

После выполнения такой процедуры над всеми сетями из табл. 5 получим ПС, представленные в табл. 7.

Сравним полученные сети с плоской сетью кластера KLAT2. Последняя имеет  $N = 64$ ,  $m = 4$ ,  $K = 31$  и  $M = 9$ . В ней 9-й коммутатор заполнен на четверть и используется также как коммутатор 2-го уровня для организации широкополосного и подключения двух резервных станций. В одном из вариантов ПС из табл. 7  $N = 73$ ,  $m = 3$ ,  $K = 31$  и  $M = 8$ , т. е. используется меньшее число карт и коммутаторов для подключения большего числа станций. Плоская сеть кластера KLAT2 может быть реализована и на  $M = 14$  значительно более дешевых коммутаторах с  $K = 23$ . У нее  $N = 73$  и  $m = 4$ .

К сожалению, для  $m = 5$  автору не удалось построить простейшую сеть 1-го ранга, которая была бы плоской. Возникает вопрос: возможно ли это теоретически? Оказывается, что для  $m \leq 6$  это всегда возможно, а для  $m > 6$  — в редких случаях. Дело в том, что ПС 1-го ранга изоморфны блок-схемам [7, 8], изученным в комбинаторике. Детальное исследование связи любых ПС и блок-схем будет проведено в отдельной работе.

В табл. 8 для  $m = 5$  приведена простейшая сеть 1-го ранга, которую удалось построить. В этой сети  $N_{1,5} = 25$ , а  $n_{1,5} = 21$ , и она не является плоской. Поэтому из нее невозможно указанным способом строить ПС с большим числом станций, а только РПС. Но об этом далее.



Таблица 6

**Освобождение одного порта со старшими номерами**

SW	P				
	0	...	21	22	23
0	1		<b>49</b>	54	50
1	2		50	<b>55</b>	51
2	3		51	<b>49</b>	<b>52</b>
3	4	...	<b>52</b>	50	53
4	5		53	51	54
5	6		54	<b>52</b>	<b>55</b>
6	7		<b>55</b>	53	<b>49</b>

Таблица 7

**Практические варианты плоских сетей**

m	K	N	M
3	23	53	7
	31	73	
	47	109	
	63	147	
	95	221	
4	23	73	13
	31	99	
	47	151	
	63	203	
	95	307	

Таблица 8

**Простейшая сеть 1-го ранга при m = 5**

SW	P				
	0	1	2	3	4
0	0	24	22	18	10
1	1	0	23	19	11
2	2	1	24	20	12
3	3	2	0	21	13
4	4	3	1	22	14
5	5	4	2	23	15
6	6	5	3	24	16
7	7	6	4	0	17
8	8	7	5	1	18
9	9	8	6	2	19
10	10	9	7	3	20
11	11	10	8	4	21
12	12	11	9	5	22
13	13	12	10	6	23
14	14	13	11	7	24
15	15	14	12	8	0
16	16	15	13	9	1
17	17	16	14	10	2
18	18	17	15	11	3
19	19	18	16	12	4
20	20	19	17	13	5
21	21	20	18	14	6
22	22	21	19	15	7
23	23	22	20	16	8
24	24	23	21	17	9

**3. РАЗРЕЖЕННЫЕ ПЛОСКИЕ СЕТИ**

Рассмотрим подробнее свойства простейшей сети 1-го ранга, представленной в табл. 8. В этой сети все группы  $PG(J)$  имеют одинаковое число станций  $n_{1,5} = 21$ , но разный их состав, и имеется  $L_{1,5} = M_{1,5} = 25$  разных групп.

Будем наращивать число станций путем увеличения числа портов в коммутаторах без увеличения их числа, т. е. строить простейшие сети 2-го, 3-го и далее рангов. Оказывается, что число разных порожденных групп при этом не увеличивается, т. е.  $L_{r,5} = L_{1,5}$ , хотя и увеличивается число станций  $n_{r,5}$  в группах. Все группы содержат равное число станций  $n_{r,5} = m_{1,5}$ . В частности, сеть 6-го ранга имеет  $N_{6,5} = 150$ ,  $K_{6,5} = 30$  и  $n_{6,5} = 126$ . Она не является плоской, но при подсоединении к коммутатору 2-го уровня становится РПС.

Простейшая сеть, будучи открытой, допускает увеличение числа станций и путем увеличения числа коммутаторов  $M > M_{r,m}$  без увеличения числа портов  $K = K_{r,m}$  в них. Это позволяет сделать базовую сеть 1-го ранга с заданным числом станций  $B_{1,m} \geq N_{1,m}$ , которая может служить строительным блоком для построения базовой сети  $r$ -го ранга на  $B_{r,m} = rB_{1,m}$  станций, которую можно превратить в РПС.

Рассмотрим это на примере построения РПС для  $N = 256$  станций с  $m = 5$  сетевыми картами на коммутаторах с  $K = 48$  портами. Для этого требуется не менее  $M = \lceil Nm/K \rceil = 27$  коммутаторов. Будем строить базовую сеть 1-го ранга с запасом по числу станций  $B_{1,5}$ , т. е. с  $B_{1,5} = M = 28$ . Простым наращиванием числа строк в табл. 8 может сделать  $n_{1,5} < 21$ . Однако путем циклического сдвига последнего столбца всегда можно восстановить значение  $n_{1,5} = 21$ . Такая базовая сеть представлена в первых шести столбцах табл. 9 — в ней  $L_{1,5} = 28$ .

Базовая сеть 10-го ранга будет содержать  $B_{10,5} = 280$  станций, подключенных к  $K = 50$  портам. В ней  $n_{10,5} = 210$  и  $L_{10,5} = L_{1,5} = 28$ . При  $K = 47$  можно уменьшить число станций до  $N = 260$ , получив  $L_{9,5} = 28$  и  $n_{9,5} = 194...197$  в разных группах. Схема подключения станций со старшими номерами к коммутаторам представлена в двух последних столбцах табл. 9 (без выделенных жирным шрифтом номеров). Номера этих станций получены путем исключения всех нечетных и части четных номеров с последующей их перенумерацией.

В эту сеть можно добавить еще две станции с выделенными жирным шрифтом номерами 260 и 261, доведя число станций до  $N = 262$ . При этом число групп увеличится до 31. Размер старых групп оказывается в диапазоне 194...199. Образуются две новые группы —  $PG(260)$  и  $PG(261)$  со 157-ю и 176-ю станциями. Последние изменения характеристик сети трудно поддаются прогнозированию.

В результате РПС содержит  $M = 29$  коммутаторов с  $K = 48$  портами каждый. Дополнительный коммутатор является коммутатором 2-го уровня для организации широковещания и соединения через два коммутатора тех станций, которые не имеют соединения через один коммутатор.

Для случаев  $m = 3$  и  $m = 4$  РПС с числом станций  $N$  и коммутаторами на  $K$  портов можно создавать аналогично. Для этого простейшая сеть 1-го ранга наращивается до базовой сети 1-го ранга путем увеличения в ней числа станций и коммутаторов до величины  $B_{1,m} = M \geq [Nm/K]$ . В такой сети  $L_{1,m} = M$ , а  $n_{1,m}$  задается его значением в простейшей сети. Затем создается базовая сеть  $r$ -го ранга с числом станций  $B_{r,m} = rB_{1,m} \geq N$ . В этой сети  $L_{r,m} = L_{1,m}$  и  $n_{r,m} = rn_{1,m}$ . Затем число узлов в ней уменьшается так, чтобы вписаться в  $K$  портов. При этом может увеличиться число групп и измениться число станций в них.

Рассмотрим это на примере создания РПС с  $m = 3$ ,  $N \geq 128$  и  $K = 24$ . Она аналогична РПС в кластере KASY0. Сначала выбираем  $B_{1,3} = M = [128 \times 3/23] = 17$ .

Базовую сеть строим так, чтобы  $n_{1,3} = 7$ . Такая сеть представлена в первых четырех столбцах табл. 10. В ней  $L_{1,3} = 17$ .

Базовая сеть 8-го ранга содержит  $B_{8,3} = 136$  станций и занимает все 24 порта во всех 17 коммутаторах. В ней  $L_{8,3} = L_{1,3} = 17$  и  $n_{8,3} = 56$ . Удалим из нее станции 120, 122, 123, 127, 128, 132 и 133. Перенумеруем оставшиеся станции и сдвинем их в 21-й и 22-й порты (два последних столбца табл. 10). Получившаяся сеть содержит  $N = 129$  станций и имеет  $L_{8,3} = 17$  и  $n_{8,3} = 52...54$ . К ней можно подсоединить и станцию с номером 129. При этом размер старых групп увеличится на 1 и появится еще одна группа  $PG(129)$ , содержащая 66 станций. Средний размер группы — 53,7. В итоге РПС содержит 18 коммутаторов и 130 станций.

Таблица 9

 Разреженные плоские сети с  $m = 5$  и  $K = 48$ 

SW	P									
	0	1	2	3	4	...	44	45	46	
0	0	27	25	21	8		232	252		
1	1	0	26	22	9		233		259	
2	2	1	27	23	10		234	253	255	
3	3	2	0	24	11		235		252	
4	4	3	1	25	12		236	254	256	
5	5	4	2	26	13		237		253	
6	6	5	3	27	14		238	<b>260</b>	<b>261</b>	
7	7	6	4	0	15		239	252	254	
8	8	7	5	1	16		240	<b>260</b>		<b>261</b>
9	9	8	6	2	17		241	253		
10	10	9	7	3	18		242		257	
11	11	10	8	4	19		243	255	254	
12	12	11	9	5	20		244		258	
13	13	12	10	6	21		245	256	255	
14	14	13	11	7	22	...	246		259	
15	15	14	12	8	23		247	<b>260</b>	256	
16	16	15	13	9	24		248			
17	17	16	14	10	25		249	255	266	
18	18	17	15	11	26		250			
19	19	18	16	12	27		251	257	256	
20	20	19	17	13	0		224		252	
21	21	20	18	14	1		225	258	257	
22	22	21	19	15	2		226		253	
23	23	22	20	16	3		227	259	258	
24	24	23	21	17	4		228	<b>261</b>	254	
25	25	24	22	18	5		229	257	259	
26	26	25	23	19	6		230	<b>261</b>		
27	27	26	24	20	7		231	258		

Таблица 10

 Разреженные плоские сети при  $m = 3$  и  $K = 24$ 

SW	P						
	0	1	2	...	20	21	22
0	0	16	14		116		128
1	1	0	15		117	119	127
2	2	1	16		118		128
3	3	2	0		102	120	119
4	4	3	1		103		
5	5	4	2		104	121	120
6	6	5	3		105	122	121
7	7	6	4		106		122
8	8	7	5	...	107	123	121
9	9	8	6		108	122	
10	10	9	7		109	124	123
11	11	10	8		110	125	124
12	12	11	9		111		125
13	13	12	10		112	126	124
14	14	13	11		113	125	
15	15	14	12		114	127	126
16	16	15	13		115	128	127

Таблица 11

 Базовые сети для  $N \geq 128$  при  $m = 3$  и  $m = 4$ 

$m$	$K$	$M$	$r$	$B_{r,m}$	$L_{r,m}$	$n_{r,m}$
3	23	17	8	136	17	56
	31	13	11	143	13	70
	47	9	16	144	9	112
	63	7	21	147	1	147
4	23	23	6	138	23	78
	31	17	8	136	17	104
	47	13	12	156	1	156



Таблица 12

**Максимальные разреженные плоские сети, набранные из плоских сетей**

$m$	$K$	$N$	$M$	$L$	$n$
3	24	159	22	3	53
	32	292	29	4	73
	48	654	43	6	109
	64	1143	64	9	127
	96	2873	92	13	221
4	32	198	25	2	99
	48	604	49	4	151
	64	1015	61	5	203
	96	2456	97	8	307

Сравним ее с РПС кластера KASY0. Последняя содержит 128 станций с 3 картами и 18 коммутаторов с 24 портами. В ней имеется 104 разные по составу станций группы, содержащие от 43 до 63 (в среднем 55) станций. На взгляд авторов, большое число разных групп является недостатком этой РПС, так как затрудняет размещение на ней параллельных задач и прогнозирование времен их выполнения.

Аналогично можно построить РПС на  $N \geq 128$  станций с  $m = 3$  картами из коммутаторов с  $K = 32$  портами. В этом случае базовая сеть 1-го ранга содержит  $B_{1,3} = M = \lceil 128 \times 3/31 \rceil = 13$  станций и коммутаторов. Она имеет  $L_{1,3} = 13$  и  $n_{1,3} = 7$ . При использовании 30 портов в коммутаторах базовая сеть 10-го ранга содержит  $B_{10,3} = 130$  станций и  $L_{10,3} = 13$  групп по  $n_{10,3} = 70$  станций. При использовании 31 порта можно сделать РПС с  $N = 133$  при  $L_{10,3} = 13$  и  $n_{10,3} = 70 \dots 71$ . Можно сделать и  $N = 134$ . При этом 13 групп содержат по 70...72 станции и появляется группа PG(133) с 91-й станцией. Напомним, что РПС использует еще 14-й коммутатор 2-го уровня.

Для распространенных коммутаторов в табл. 11 приведены базовые сети наибольших рангов при одном коммутаторе 2-го уровня. Для них выбирается минимальное число станций  $N \geq 128$ . Дополнительно в табл. 11 приводится ранг сети, число групп и их размер. Отметим, что РПС с  $m = 3$ ,  $K = 63$  и  $m = 4$ ,  $K = 47$  превратились в ПС (сравни с табл. 7).

Основной недостаток РПС, построенных таким способом, состоит в том, *не любая* пара станций в каждой группе имеет соединение через один коммутатор. Более того, только небольшая часть станций в каждой группе имеют такое соединение — только те, которые подсоединены к одному и тому же коммутатору.

Для случаев  $m = 3$  и  $m = 4$  существует еще один способ построения РПС, который свободен от указанного недостатка. Разреженную ПС можно набирать из соответствующих ПС, представленных в табл. 7, объединенных коммутатором 2-го уровня. Например, из  $M = 22$  коммутаторов с  $K = 24$  портами можно собрать РПС на  $N = 159$  станций с  $m = 3$  картами. В ней имеется

$L = 3$  группы по  $n = 53$  станции, в которых возможно соединение через один коммутатор *любой* пар станций. Аналогично, из  $M = 15$  коммутаторов с  $K = 32$  портами можно собрать РПС на  $N = 146$  станций с  $m = 3$  картами. В ней имеется  $L = 2$  группы по  $n = 73$  станций. Максимальные конфигурации подобных сетей с одним коммутатором 2-го уровня представлены в табл. 12.

Очевидный недостаток этих РПС состоит в малой пропускной способности и больших задержках для станций, которые соединяются через единственный коммутатор 2-го уровня. Этот недостаток можно преодолеть, если построить ПС 2-го уровня. В этой сети коммутаторы 1-го уровня должны выделять  $q > 1$  портов для соединения с  $Q > 1$  коммутаторами 2-го уровня, и должно существовать соединение через один коммутатор 2-го уровня между любой парой выделенных портов коммутаторов 1-го уровня.

### ЗАКЛЮЧЕНИЕ

Предложен регулярный способ построения плоских и разреженных плоских сетей для построения дешевых кластерных суперкомпьютеров из персональных компьютеров на основе коммутируемых сетей (Fast)Ethernet и программного обеспечения Beowulf. Он обеспечивает более высокие характеристики плоских и разреженных плоских сетей, чем способ стохастической оптимизации конфигурации таких сетей на основе генетических алгоритмов.

Для повышения пропускной способности разреженных плоских сетей предложено создавать их с использованием плоских сетей 1-го и 2-го уровней. Поставлена задача построения простейших плоских сетей для станций с 5-ю (и, возможно, с 6-ю) сетевыми картами.

### ЛИТЕРАТУРА

1. Adams J. C., Laverell W. D., Ryken M. A. MBH'99: A Beowulf cluster capstone project / <http://www.calvin.edu/~adams/html/interests/professional/publications/MBH99.pdf>.
2. Deitz H. G. and Mattox T. I. KLAT2's flat neighborhood network / <http://aggregate.org/FNN/>.
3. Deitz H. G. and Mattox T. I. Compiler techniques for flat neighborhood networks / <http://aggregate.org/FNN/>.
4. Алленов А. В., Подлазов В. С., Стецюра Г. Г. Пропускная способность набора кольцевых каналов. I. Класс наборов колец. Наборы с простыми узлами // Автоматика и телемеханика. — 1996. — № 3. — С. 135—144.
5. Подлазов В. С., Подлазова А. В. Обеспечение наращиваемости отказоустойчивых многопроцессорных систем с общей памятью с использованием многокольцевых некоммутируемых сетей связи с неоднородными узлами // Труды Ин-та пробл. управл. РАН. — 2002. — Т. XVIII. — С. 164—181.
6. <http://aggregate.org/KASY0/>.
7. Finn A. M., Decker R. O. A network architecture for radar signal processing // AIAA/IEEE 8-th digital avionic systems conference / San Jose. Calif., Oct. 1988. — P. 614—621.
8. Холл М. Комбинаторика — М.: Мир, 1970. — Гл. 10.

☎ (495) 334-78-31

E-mail: podlazov@ipu.ru

