

УДК 519.87:519.6:624.9

О некоторых характеристиках циркулянтных и тороидальных структур вычислительных систем

Э. А. Монахова, О. Г. Монахов

В работе даны краткий обзор перспективных топологий высокопроизводительных кластерных параллельных систем, использующих тороидальные и циркулянтные структуры сетей связи размерностей больших двух, и сравнительный анализ их структурных характеристик. Показано преимущество циркулянтных структур по таким структурным характеристикам, как диаметр, средний диаметр и ширина бисекции, при соизмеримых затратах на число узлов и число связей системы.

Ключевые слова: параллельные кластерные системы, тороидальные структуры, циркулянтные сети, диаметр, средний диаметр, ширина бисекции.

1. Введение

Циркулянтные сети широко изучаются в теории графов и дискретной математике, а также играют важную роль в разнообразных приложениях, в том числе теории кодирования, параллельных вычислениях, при построении топологий сетей и мультикластерных систем [1, 2, 3, 4, 5, 6, 7, 8].

Пусть N и $s_1, s_2, \dots, s_k$ – целые числа такие, что $1 \leq s_1 < s_2 < \dots < s_k < N$. Неориентированный граф $C(N; S)$, где $S = \{s_1, s_2, \dots, s_k\}$, с множествами вершин $V = \{0, 1, \dots, N-1\}$ и рёбер $E = \{(i, j) : |i - j| \equiv s_m \pmod{N}, m = \overline{1, k}\}$, называется *циркулянтным*, числа из множества S – *образующими*, k – *размерностью*, N – *порядком* графа. Обзоры результатов по исследованию циркулянтных графов можно найти в [2, 3, 4]. Следует отметить, что открытые в литературе семейства двумерных и трёхмерных циркулянтных сетей (а в последнее время и более высоких размерностей [5]) также успешно применяются в теории кодирования при построении совершенных кодов, исправляющих ошибки.

Дадим формальное определение другой популярной топологии: $T = (V, E)$ есть k -мерный тор порядка $N = \prod_{i=1}^k l_i$, если $V = \{(v_1, v_2, \dots, v_k) : 0 \leq v_i < l_i\}$, $E = \{(u, v) : \exists j \text{ такое, что для } \forall (i \neq j) v_i = u_i \text{ и } v_j \equiv u_j \pm 1 \pmod{l_j}\}$. Обозначим такой тор как $T_{l_1, l_2, \dots, l_k}^{(k)}$.

Определение эффективной топологии сети связей при проектировании высокопроизводительных параллельных систем является одним из основных вопросов, который приобретает особую важность при выходе на размеры систем $10^6 - 10^9$ процессоров (узлов).

Структурными характеристиками графа, по которым в первую очередь сравниваются различные топологии систем, являются диаметр и среднее расстояние (оценивают, соответственно, максимальную и среднюю структурные задержки в сети, а также время выполнения коллективных обменов в системе). *Диаметром* графа $G = (V, E)$ называется величина

$$D(G) = \max_{i, j \in V} d(i, j),$$

где $d(i, j)$ – длина кратчайшего пути из вершины i в вершину j . Средним расстоянием (*средним диаметром*) графа порядка N называется $d_{av} = \frac{1}{N(N-1)} \sum_{i,j \in V, i \neq j} d(i, j)$.

Ещё одним типичным показателем для оценки и сравнения топологий параллельных архитектур (кластерных систем) является *ширина бисекции* графа BW (bisection width) – минимальное число рёбер, которые должны быть удалены, чтобы граф $G = (V, E)$ разделился на две равные части G^1 и G^2 такие, что $|V(G^1)| = \lfloor N/2 \rfloor$ и $|V(G^2)| = \lceil N/2 \rceil$, где $N = |V(G)|$. Ширина полосы пропускания BBW (bisection bandwidth) оценивает количество данных, которые могут быть переданы между этими частями графа. Вторая характеристика зависит от первой, поэтому большинство работ посвящено исследованию ширины бисекции графа. Проблема нахождения ширины бисекции графа является NP -полной, но нижние и верхние границы известны для большинства регулярных топологий сетей связи (гридов, гиперкубов, торов).

Цель настоящей работы – обзор перспективных высокопроизводительных кластерных параллельных систем, использующих тороидальные и циркулянтные топологии размерностей больших двух, и сравнительный анализ их структурных характеристик.

2. Перспективные кластерные системы больших размерностей

При проектировании серии кластерных систем Swiss-Tx [8] и выборе топологии сети связи для них проведён [9] сравнительный анализ трёх перспективных классов структур: торов, циркулянтных графов и fat-деревьев [10]. Гиперкубы, несмотря на малый диаметр и хорошие математические свойства, не рассматривались вследствие существенного роста числа связей в каждой вершине при увеличении размерности и, следовательно, низкой масштабируемости (существованию только для отдельных редких значений порядка графа). Классы структур сравнивались по трём показателям: диаметру, среднему диаметру и ширине бисекции (ширине полосы пропускания) при одинаковых порядках и степенях графов. Для анализа взяты циркулянты вида $C(N; 1, s_2, \dots, s_k)$, размерностей $k = 2, 3, 4, 5$, исходя из их возможной практической реализации. Для определения диаметра и среднего диаметра циркулянтов использовались аппроксимационные формулы. Кроме того, анализ топологий проводился с использованием crossbar switches, как это имеет место в реальных кластерных системах. Результаты анализа показали, что из сравниваемых структур торы всегда имеют самые плохие диаметр, средний диаметр и ширину бисекции. Недостаток fat-деревьев при хороших диаметре и ширине бисекции – плохая наращиваемость, тогда как циркулянты существуют при любом числе вершин и всегда можно компенсировать несколько больший диаметр и меньшую ширину бисекции по сравнению с fat-деревьями увеличением степени графа. Топология циркулянтного графа позволяет адаптировать производительность системы к потребностям пользователя и получить при более низкой стоимости производительность, эквивалентную или лучшую, чем у fat-деревьев. Более того, топология fat-деревьев не может полностью извлечь выгоду из увеличивающегося размера технологии crossbar switches, поскольку степень fat-дерева должна соответствовать половине размера $n \times n$ -crossbar, и число процессоров должно быть степенью $n/2$.

Реализация и эксплуатация кластерной системы Swiss-T1 с циркулянтной сетью связи показала сетевые характеристики достаточно мощные, чтобы обеспечить приложения, требующие высокой пропускной способности и низкой задержки.

Суперкомпьютерные системы Blue Gene/L, Blue Gene/P [11] с числом вычислительных узлов до 32768 используют топологию трёхмерного тора в качестве интерконнекта с высокими показателями пропускной способности и масштабируемости и применяются для решения многих научных проблем и массово-параллельных задач в генетике. Суперкомпьютер Blue Gene/Q использует уже топологию 5D-тора [12]. Японская фирма Fujitsu реализовала

К-computer [13] с оригинальной топологией на основе 3D-тора, каждый узел которого иерархически включает структуру из 12 процессоров, с общим числом до 80000 узлов.

Суперкомпьютеры СКИФ ряда 4 [14] – отечественные кластерные системы, использующие топологию 3D-тора. Системная сеть СКИФа реализована в виде трёхмерного тора с числом вершин до $N = 6400 = 16 \times 16 \times 25$. Предусмотрено расширение количества вычислительных узлов в каждой стойке с 256 до 736. В будущем планируется в качестве системной сети использование nD-торов более высоких размерностей. На повестке дня стоят проблемы создания перспективного интерконнекта: эффективная топология с масштабируемостью до $10^6 - 10^9$, эффективной маршрутизацией, низкой задержкой, высокими пропускной способностью и темпом выдачи сообщений.

Четырёхмерные циркулянтные графы использованы в качестве базовых элементов топологии при построении мультипроцессорной кластерной системы [7]. Топология кластерной системы, построенной в университете г. Манилы, основана на 16 копиях циркулянтного графа вида $C(16; 1, 4, 7, 8)$. В целом система с использованием материнских плат с четырьмя входами/выходами имеет 1024 процессора. Таким образом, это – пример реализованной кластерной системы с циркулянтной топологией размерности, большей трёх. Четырёхмерный циркулянт $C(16; 1, 4, 7, 8)$ в данном случае имеет степень вершины $v = 7$, диаметр $D = 2$ и значение ширины бисекции $BW = 16$.

3. Сравнительный анализ структурных характеристик систем

Диаметр k -мерного, $k \geq 2$, тора с числом вершин $N = \prod_{i=1}^k l_i$ равен $D = \sum_{i=1}^k \lfloor \frac{l_i}{2} \rfloor$ [15]. Известно, что торы дают меньший диаметр и лучшую производительность, если число вершин в каждом направлении одинаково, то есть $l_i = l$ для всех $i = 1, 2, \dots, k$. В этом случае диаметр тора равен $D = k \lfloor l/2 \rfloor$, а средний диаметр – $d_{av} = \lfloor \frac{l^2}{4} \rfloor k/l$. Точное значение ширины бисекции для k -мерных торов общего вида $T_{l_1, l_2, \dots, l_k}^{(k)}$, которые имеют l_i вершин вдоль размерности i для $i = 1, 2, \dots, k$ и где $l_1 \geq l_2 \geq \dots \geq l_k$, найдено недавно в [16]:

$$BW(T_{l_1, l_2, \dots, l_k}^{(k)}) = 2 \sum_{i=1}^{\alpha} \prod_{j=i+1}^k l_j,$$

где α – наименьший индекс, для которого l_i – чётное число ($\alpha = k$, если такого индекса нет).

Для многомерных циркулянтных графов в [17] найдена следующая верхняя оценка ширины бисекции: для любого циркулянтного графа степени $2k$ с N вершинами и множеством образующих $S = \{s_1, s_2, \dots, s_k\}$, где $1 \leq s_1 < s_2 < \dots < s_k < N/2$,

$$BW \leq 2 \sum_{i=1}^k s_i.$$

Учитывая зависимость этой оценки от суммы значений образующих, можно её определить как $BW \leq 2 \min_{T \sim S} \sum_{i=1}^k t_i$, где $\sim$ – отношение эквивалентности между двумя множествами образующих $T = \{t_1, t_2, \dots, t_k\}$ и S .

В качестве перспективной топологии для суперкомпьютеров с трёхмерной сетью связи и числом процессоров $N > 4 \times 10^3$ могут быть рекомендованы оптимальные циркулянтные сети размерности три с максимальным числом вершин для любого диаметра $D \geq 1$, описываемые аналитически [18]:

$$\{C(N(p); 1, s_2(p), s_3(p)) | p = 2 \lfloor D/3 \rfloor\},$$

где $N = 4p^3 + 4p^2 + 3p + 1$, $s_2 = 2p^2 + p$, $s_3 = 2p^2 + 3p + 2$, если $D \equiv 0 \pmod{3}$;
 $N = 4p^3 + 12p^2 + 15p + 7$, $s_2 = 2p^2 + 3p + 2$, $s_3 = 2p^2 + 5p + 4$, если $D \equiv 1 \pmod{3}$;
 $N = 4p^3 + 20p^2 + 35p + 21$, $s_2 = 2p^2 + 5p + 4$, $s_3 = 2p^2 + 7p + 6$, если $D \equiv 2 \pmod{3}$.

Для них также имеются эффективные бестабличные алгоритмы маршрутизации, обеспечивающие минимумы времени выполнения и нагрузки сообщений в сети и допускающие простые модификации, учитывающие отказы процессоров и линий связи [18].

Важно отметить, что циркулянтные графы с максимальным (или близким к нему) порядком для заданных размерности и диаметра, взятые в качестве сетей связи мультипроцессорных систем, являются наиболее плотными графами и, соответственно, имеют высокие отказоустойчивость, пропускную способность и темп выдачи сообщений, минимальную задержку и максимальные связность и надёжность [1, 3]. Поэтому далее будем проводить сравнение трёхмерных и четырёхмерных тороидальных структур с семействами оптимальных циркулянтов.

Для трёхмерных оптимальных циркулянтных сетей верхняя граница ширины бисекции может быть получена аналитически из приведённой выше верхней оценки BW (как функция от $D > 1$, $p = 2\lfloor D/3 \rfloor$):

$$BW \leq \begin{cases} 8p^2 + 8p + 6 & \text{при } D \equiv 0 \pmod{3}, \\ 8p^2 + 16p + 14 & \text{при } D \equiv 1 \pmod{3}, \\ 8p^2 + 24p + 22 & \text{при } D \equiv 2 \pmod{3}. \end{cases}$$

Были проведены численные эксперименты с целью получения точных значений ширины бисекции для рассматриваемых оптимальных циркулянтов. Программа случайного поиска минимума ширины бисекции на достаточно представительной выборке (10^6 испытаний при числе вершин $N = 21, 55, 117$) не нашла меньших значений, чем представленные выше. Таким образом, полученные выше аналитические значения ширины бисекции являются с большой вероятностью её точными значениями.

В табл. 1 проведено сравнение топологий циркулянтов и 3D-торов по следующим показателям: диаметру (D), среднему диаметру (d_{av}) и ширине бисекции (BW). Порядки циркулянтов в табл. 1 (а также в табл. 2) являются наиболее близкими к числу вершин соответствующих торов. Значения среднего диаметра циркулянтов определены с помощью программы анализа структурных характеристик.

Т а б л и ц а 1. Сравнение топологий 3D-торов и циркулянтов размерности три

3D-torus	$N(torus)$	D	d_{av}	BW	s_2, s_3	$N(circ)$	D	d_{av}	BW
$16 \times 16 \times 16$	4096	24	12	512	210, 232	4431	15	11.27	886
$16 \times 16 \times 25$	6400	28	14,25	544	254, 276	6371	17	12.73	1062
$24 \times 24 \times 24$	13824	36	18	1152	436, 466	13545	22	16.39	1806
$32 \times 32 \times 32$	32768	48	24	2048	820, 862	33661	30	22.22	3366
$48 \times 48 \times 48$	110592	72	36	4608	1830, 1892	111691	45	33.17	7446
$64 \times 64 \times 64$	262144	96	48	8192	3240, 3322	262521	60	44.11	13126
$96 \times 96 \times 96$	884736	144	72	18432	7382, 7504	908107	91	66.71	29774
$128 \times 128 \times 128$	2097152	192	96	32768	12880, 13042	2073841	120	87.86	51846
$136 \times 136 \times 136$	2515456	204	102	36992	14878, 15052	2574067	129	94.42	59862

Для четырёхмерных кластерных систем в качестве топологии могут быть рекомендованы (как увеличивающие ширину бисекции и улучшающие соотношение N/D) семейства циркулянтов размерности четыре с наилучшим известным в литературе соотношением N/D , $D \geq 2$, следующего аналитического описания [19]:

$$\{C(N(p); 1, s_2(p), s_3(p), s_4(p)) | p = D/2 \geq 1\},$$

где $N = 8p^4 + 8p^3 + 10p^2 + 4p + 1$, $s_2 = 2p + 1$, $s_3 = 4p^3 + 2p^2 + 2p$, $s_4 = 4p^3 + 6p^2 + 6p + 2$ при чётном D ; $s_2 = 2p$, $s_3 = 4p^3 + 3p$, $s_4 = 4p^3 + 4p^2 + 3p + 1$ при нечётном D .

В табл. 2 даны значения трёх основных структурных показателей для четырёхмерных торов и циркулянтов размерности четыре. Значения ширины бисекции для циркулянтов вычислялись как удвоенная сумма соответствующих образующих.

Т а б л и ц а 2. Сравнение топологий 4D-торов и циркулянтов размерности четыре

4D-torus	$N(\text{torus})$	D	d_{av}	BW	s_2, s_3, s_4	$N(\text{circ})$	D	d_{av}	BW
$8 \times 8 \times 8 \times 8$	4096	16	8	1024	9, 378, 460	4231	9	7.2	1696
$12 \times 12 \times 12 \times 12$	20736	24	12	3456	15, 1484, 1710	22471	14	11.06	6420
$16 \times 16 \times 16 \times 16$	65536	32	16	8192	19, 3096, 3458	59167	18	14.14	13148
$20 \times 20 \times 20 \times 20$	160000	40	20	16000	23, 6118, 6648	153437	23	17.98	25580
$24 \times 24 \times 24 \times 24$	331776	48	24	27648	29, 11396, 12238	331297	28	21.82	47328
$32 \times 32 \times 32 \times 32$	1048576	64	32	65536	39, 28196, 29718	1101127	38	29.5	115908
$40 \times 40 \times 40 \times 40$	2560000	80	40	128000	47, 51982, 54192	2549281	47	36.40	212444

На рис.1 представлены графики зависимости диаметра D и ширины бисекции BW от числа вершин N для рассматриваемых циркулянтов и торов.

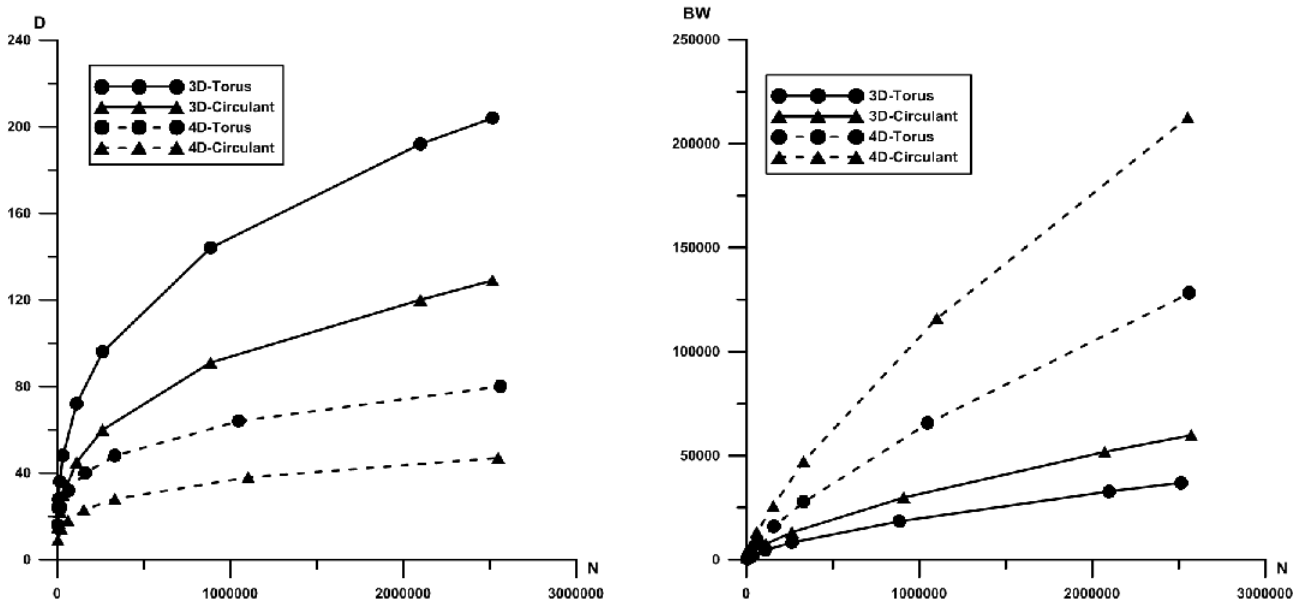


Рис. 1. Графики зависимости диаметра D и ширины бисекции BW от числа вершин N для рассматриваемых циркулянтов и торов

Отметим, что для топологии кластерной системы из [7] наилучшее соотношение N/D при диаметре $D = 2$ и размерности $k = 4$ имеет максимально возможный граф с числом вершин $N = 35$ и образующими $S = \{1, 6, 7, 10\}$ [19]. Образующие $S = \{1, 2, 7, 12\}$ для порядка $N = 16$ дают возможность наращивать число вершин базового графа от 16 до 29 ценой минимальной перестройки структуры, сохраняя при этом диаметр два и значения образующих. Численные эксперименты с целью получения точных значений ширины бисекции дали также улучшение этого показателя до $BW = 28$ для графа $C(16; 1, 2, 4, 7)$, который эквивалентен графу $C(16; 1, 2, 7, 12)$.

4. Заключение

Улучшение рассмотренных структурных показателей у циркулянтных сетей указанных размерностей по сравнению с торами (более чем в 1.5 раза для диаметра и ширины бисекции

и на 10% для среднего диаметра) позволяет уменьшить задержки при передаче информации, сократить время исполнения заданий, увеличить показатели надёжности и живучести системы, и, соответственно, при соизмеримых затратах на число узлов и число связей получить большую производительность и экономическую эффективность.

Литература

1. Монахов О.Г., Монахова Э.А. Параллельные системы с распределённой памятью: структуры и организация взаимодействий. Новосибирск: Изд-во СО РАН, 2000. 242 с.
2. Монахова Э.А. Структурные и коммуникативные свойства циркулянтных сетей // Прикладная дискретная математика. 2011. №3(13). С. 92-115.
3. Bermond J.-C., Comellas F., Hsu D.F. Distributed loop computer networks: a survey // J. Parallel Distributed Comput. 1995. V. 24. P. 2-10.
4. Hwang F.K. A survey on multi-loop networks // Theoretical Computer Science. 2003. V. 299. P. 107-121.
5. Martinez C., Beivide R., and Gabidulin E. M. Perfect codes from Cayley graphs over Lipschitz integers // IEEE Transactions on Information Theory. 2009. V. 55, №8. P. 3552-3562.
6. Нестеренко Б. Б., Новотарский М. А. Клеточные нейронные сети на циркулянтных графах // Искусственный интеллект. 2009. № 3. С. 132-138.
7. Muga II F.P., Saldana and Yu W.E.S. Building Graph-Based Symmetric Cluster // NECTEC Technical Journal. 2001. V. 11, №9. P. 195-199.
8. Gruber R., Volgers P., De Vita A., Stengel M. and the Swiss-Tx Team. Commodity computing results from the Swiss-Tx project //Preprint of Ecole Polytechnique Federal, CH-1015, Lausanne, Switzerland. 2001, 13 p.
9. Kuonen P., Gruber R. Parallel computer architectures for commodity computing and the Swiss-T1 machine // EPFL Supercomputing review. 1999. №1. P. 3-11.
10. Leiserson Ch. E. Fat-trees: universal networks for hardware-efficient supercomputing. //IEEE Transactions on Computers, 1985. V. 34 , № 10. P. 892-901.
11. Adiga N.R. et. al. Blue Gene/L torus interconnection network // IBM J. Res. Dev. 2005. V. 49, № 2/3. P. 24-35
12. Chen D., et. al. The IBM Blue Gene/Q interconnection network and message unit. // Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis (SC11). Seattle, Washington, USA, November 12-18, 2011. P.26.
13. Innovative «6-Dimensional Mesh/Torus» Topology Network Technology. Fujitsu Limited. [Электронный ресурс]. URL: <http://www.fujitsu.com/global/about/tech/k/whatis/network/> (дата обращения: 20.05.2013).
14. Абрамов С.М., Заднепровский В.Ф., Лилитко Е.П. Суперкомпьютеры СКИФ ряда 4 // Информационные технологии и вычислительные системы. 2012. № 1. С. 3-16.

15. Воробьев В.А. Простейшие структуры однородных вычислительных систем // Вычислительные системы. Вопросы теории и построения вычислительных систем. Новосибирск. 1974. Вып.60. С. 35-49.
16. Aroca J.A. and Anta A.F. Bisection (Band)Width of Product Networks with Application to Data Centers.// Theory and Applications of Models of Computation. Lecture Notes in Computer Science. 2012. V. 7287. P. 461-472.
17. Mans B. and Shparlinski I. Bisecting and Gossiping in Circulant Graphs// 6th Latin American Theoretical Informatics Conf., Buenos Aires, 2004. LNCS. 2004. V. 2976. P. 589-598.
18. Монахова Э. А. Циркулянтные сети связи вычислительных систем: структуры и обмены // Материалы 5-го Международного научно-практического семинара «Высокопроизводительные параллельные вычисления на кластерных системах». Нижний Новгород. 2005. С.175-181.
19. Монахова Э. А. Новая достижимая нижняя оценка числа вершин в циркулянтных сетях размерности четыре // Дискретный анализ и исследование операций. 2013. Т. 20. №1. С.37-44.

*Статья поступила в редакцию 20.05.2013;
переработанный вариант – 30.05.2013.*

Монахова Эмилия Анатольевна

канд. техн. наук, доцент, старший научный сотрудник, Институт вычислительной математики и математической геофизики СО РАН (630090, Новосибирск, Лаврентьева пр., 6), тел. (383)3011489, e-mail: emilia@rav.sccc.ru

Монахов Олег Геннадьевич

канд. техн. наук, ведущий научный сотрудник, Институт вычислительной математики и математической геофизики СО РАН (630090, Новосибирск, Лаврентьева пр., 6), тел. (383)3011489, e-mail: monakhov@rav.sccc.ru

Some characteristics of circulant and toroidal structures of computer systems

Monakhova E.A., Monakhov O.G.

The paper presents a brief overview of promising high-performance network topologies of parallel systems. For toroidal structures and circulant networks with dimension larger than two, a comparative analysis of their structural characteristics is given. The advantage of circulant networks for structural characteristics such as diameter, average diameter and bisection width, with comparable costs to the number of nodes and the number of connections is demonstrated.

Keywords: parallel cluster systems, toroidal structures, circulant networks, diameter, average diameter, bisection width.