УДК 681.321

М.Ф. КАРАВАЙ, П.П. ПАРХОМЕНКО, В.С. ПОДЛАЗОВ

Институт проблем управления им. В.А. Трапезникова РАН, Россия

К НОВОЙ ТОПОЛОГИИ ОТКАЗОУСТОЙЧИВЫХ КЛАСТЕРОВ И ЛОКАЛЬНЫХ УПРАВЛЯЮЩИХ СЕТЕЙ

Малоизвестная в инженерно-технических кругах математическая комбинаторная конструкция *симметричных уравновешенных блок-схем* (block-designs) содержит большие возможности для проектирования отказоустойчивых неоднородных кластеров и локальных сетей для сбора информации и управления. При надлежащей интерпретации, блок-схемы можно рассматривать как квазиполносвязный граф, вершины которого соединены не по принципу ”точка-точка”, а через достаточно простой переключатель, практически не вносящий дополнительной задержки при прохождении сигналов. При этом число каналов связи и портов n-узловой сети уменьшается в $\sqrt{n}$ раз по сравнению с полнодоступным графом.

модулярная арифметика, система остаточных классов, отказоустойчивость, надежность, производительность

Введение

Глубинной целью проектирования отказоустойчивых систем, не всегда явно провозглашаемой и даже не всегда чётко осознаваемой, выступает такое проектирование (синтез), в котором “целевая” структура является инвариантом. Под “ целевой ” понимается исходная неизбыточная система. Инвариантность предполагает неизменность основных свойств системы при её преобразованиях, например, под действием неисправностей, при масштабировании или при введении избыточности. В частности, задача может быть поставлена как сохранение логической структуры исходной задачи при отказах. Это требует нахождения минимального избыточного графа, в который вкладывается целевой граф, и разработки алгоритма реконфигурации для нахождения графа, изоморфного целевому [1]. Исследования в этом направлении показали, что для ряда популярных однородных структур, таких как кольца, хордальные кольца, решётки и некоторые другие могут быть найдены минимальные решения для синтеза 1-отказоустойчивых структур. Однако, с ростом степени отказоустойчивости (>1) и с ростом числа узлов (вершин) в структуре минимальным решением становится полносвязный граф, т.е. очень дорогое, но тривиальное решение. И улучшить здесь ничего нельзя, т.к. решение минимальное. Проблема ещё более усложняется, рассматривая неоднородные системы. Здесь не проглядывается иного выхода, чем переход к структуре системы в виде полного графа. Практические реализации подтверждают этот вывод. Для небольшого числа вершин (до 8, 16) может быть создана полносвязная параллельная система. Для бóльшего числа узлов выход находится в переходе к шинной архитектуре, в которой полносвязность реализуется последовательно во времени.

Понятно, что поддержание высокой производительности требует очень высокой рабочей частоты на шине для эмуляции параллельной полносвязности. Итак, с одной стороны, полносвязность идейно подходит для организации отказоустойчивости, наращиваемости и работы в неоднородной приборной среде, но, с другой стороны, совершенно неприемлема из соображений стоимости. В некотором смысле промежуточное положение занимает архитектура сети Space-wire [3]. Но в ней повышение производительности опирается на увеличение пропускной способности канала, а не на топологию “быстрой” сети.

Возникает вопрос, можно ли спроектировать сетевую топологию, обладающую лучшими свойствами из ранее применявшихся топологий?

Рассмотренные характеристики: *стоимость, производительность, отказоустойчивость, расширяемость, реконфигурируемость, контролеспособность, простота внесения избыточности* взаимозависимы и взаимно противоречивы. Конечно, всё было бы проще, если бы отказоустойчивость всегда вела к снижению стоимости, лучшей производительности, масштабируемости и др. характеристик. Но это объективно не так.

Возможно ли найти подходящую *базовую* математическую конструкцию, которая, как это видно из изложенного выше, была бы по положительным свойствам близка к полному графу, но значительно проще и дешевле? При положительном ответе открылась бы возможность исследования проблемы, имеющей прямое отношение к отказоустойчивости ***сетевых и кластерных систем***, в том числе неоднородных. В рамках поиска структурных инвариантов сложных систем задача звучит как “сохранение связности логических абонентов кластера или сети при возможных отказах”.

Двудольные графы и уравновешенные блок-схемы

Такая математическая конструкция существует. В частности, это семейство двудольных (состоящих из двух непересекающихся подмножеств X и Y вершин) графов, обладающих тем свойством, что каждая вершина из X (из Y) связана с каждой из остальных вершин из X (из Y) единственным путем длины 2, проходящим через некоторую вершину из Y (из X).

Здесь будут сформулированы преимущества подобных сетей перед известными последовательными и параллельными сетями, показано, что такие сети можно рассматривать как “квазиполные”.

Прежде всего, впервые было замечено, что *двудольные графы* (bipartite graphs), уравновешенные блок-схемы (block-designs) и коммутационные сети– это не разрозненные понятия, а “родные братья”. Это оказалось важным, поскольку дало в руки сильный математический инструмент – теорию уравновешенных блок-схем в математической комбинаторике [2]. Поскольку этот инструмент довольно сложный, то он практически незнаком инженерам. Математики, со своей стороны, никак не связывали его с такими приложениями как *сети*. Увидев это, удалось сформулировать задачу в виде проектирования *высокопроизводительных отказоустойчивых* систем коммутации в кластерах и локальных сетях [4].

Главным объектом исследований в [2] являются неполные уравновешенные блок-схемы, которые представляют размещение $\boldsymbol{v}$ элементов по $\boldsymbol{b}$ блокам таким образом, что каждый блок содержит по $\boldsymbol{k}$ различных элементов, каждый элемент принадлежит $\boldsymbol{r}$ различным блокам, а каждая пара элементов появляется в $\boldsymbol{\lambda}$ блоках. Если “обычный” граф задаётся инцидентностями к рёбрам своих вершин, то комбинаторная блок-схема представляет собой более сложный комплекс инцидентностей, задаваемый на некотором исходном множестве, и существование этого комплекса часто далеко не очевидно.

Здесь рассматриваются *симметричные* блок-схемы, у которых $v = b = n$ - число вершин подмножеств X и Y, $k = r = s$ - степень вершин двудольного графа и $\lambda = 1$.

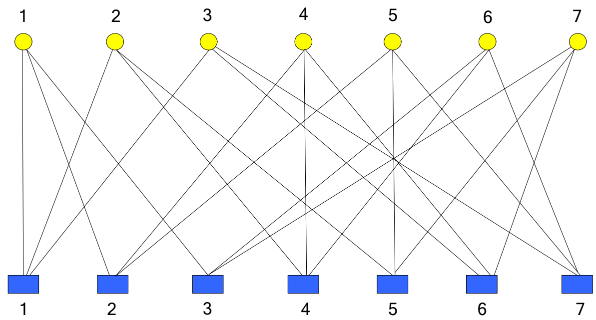

Рис. 1. Двудольный граф

На рис.1 дан пример двудольного графа с X=Y=7, ***s***=3. “Сетевое” представление блок-схемы 13х13, ***s***=4 приведено на рис. 2. Число *n* элементов подмножеств X и Y должно удовлетворять равенству:

$$n = s\,(s - 1) + 1. \qquad (*)$$

Поэтому симметричный двудольный неизбыточный граф (как симметричная блок-схема [2]) может иметь число вершин подмножеств X и Y (число элементов и блоков блок-схемы) только из ряда значений, начало которого приведено в табл. 1. Такой граф логично назвать *квазиполносвязным*.

Таблица 1

s	2	3	4	5	6	7	8	9	10	11	…
n	3	7	13	21	31	43	57	73	91	111	…

Следует заметить, что условие (*) необходимое, но не для всех *n* и *s* достаточное. В частности, блок-схема для *s*=7, *n*=43 нереализуема [2].

Техническая суть

Если множество абонентов соединить между собой сетью в виде полного графа, т.е. “каждый непосредственно с каждым”, то решаются почти все проблемы, кроме одной важнейшей – сложности. Например, для локальной сети на 32 абонента каждая вершина (абонент) должен иметь 31 порт и соответствующий шлейф соединений, что почти немыслимо для такой небольшой задачи. Но при этом производительность будет максимальной, наращиваемость – однородной, программное обеспечение межабонентной коммутации – простейшим. Однако упомянутая сложность всё это перечёркивает.

Сеть в виде реализации *блок-схемы* позволяет изящно решить все эти проблемы математически, так сказать, по максимально возможному варианту, т.е. оптимально. Математика говорит, что такие сети *существуют*. Разработка алгоритмов синтеза рассматриваемых блок-схем представляет серьёзную трудность. За небольшим исключением (где *s*-1 = *p* - простое число) для каждого *s* приходится изобретать алгоритм индивидуально. Величина *s* определяет число абонентов в кластере, однородно соединённых друг с другом. В табл. 1 оно определяется по разности значений *n* между соседними клетками (кроме пары s=7, n=43). В пределах каждого кластера вопросы наращиваемости, сохранения быстродействия, отказоустойчивости и связности решаются без особых проблем.

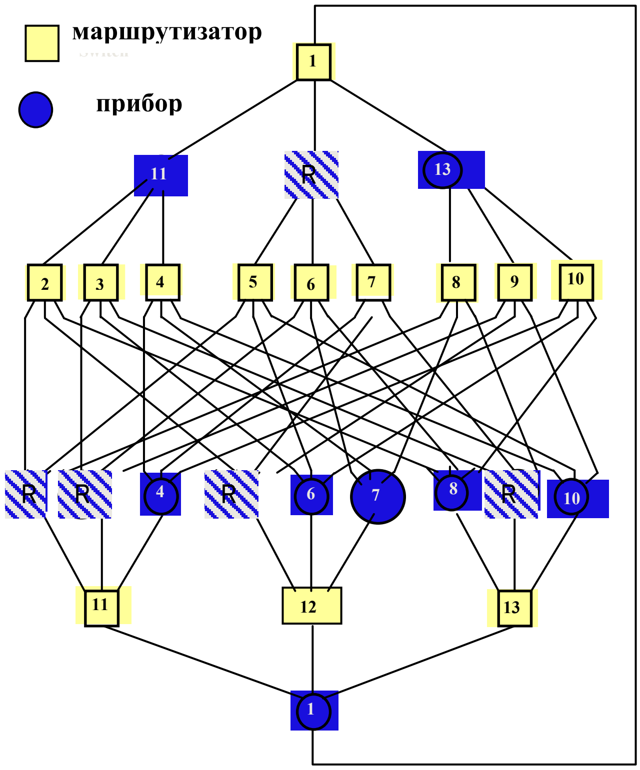

Рис. 2. Сетевая реализация блок-схемы 13х13

Пока ещё нет непрерывного заполнения таблицы 1 по всем значениям *s>9,* но работа в этом направлении ведётся. Каждое решение представляет существенный интерес для локальных промышленных сетей гетерогенных абонентов. При этом число портов в каждой вершине растёт *как корень квадратный* из числа абонентов, а не пропорционально, как в “старой” полносвязной сети. При этом легко вводятся избыточные абоненты, необходимые для отказоустойчивости, и выполняется требование сохранения такого *инварианта сети как связность каж-*

дой пары абонентов, несмотря на возможное появление отказов в системе.

Популярные топологии кластеров и коммуникационных сетей

Уже отмечалось, что многие популярные топологии коммуникационных сетей обладают такими недостатками как невозможность удобного наращивания, повышения степени отказоустойчивости выше значения 1, удобного введения резерва, использования в гетерогенных структурах и т.п.

Структура 5x4 может быть достаточно экономно сделана 1-отказоустойчивой. Однако наращивание на любое число узлов становится проблематичным, поскольку требует изменения структурных решений. Столь же проблематично построение 2-х или 3-х отказоустойчивых структур для этой решётки, т.к. топология быстро приближается к полносвязному графу.

Подобные проблемы характерны и для других топологий, в частности, для приведённых на рис. 3.

Проще всех указанные проблемы решаются для шины, поскольку она эмулирует во времени полносвязный граф. Но и производительность её - минимальная по сравнению с другими топологиями. Многокаскадная коммутационная сеть типа Ω-сети на рис. 4, как и многомерный куб, тоже обладают плохой наращиваемостью и неэкономной отказоустойчивостью.

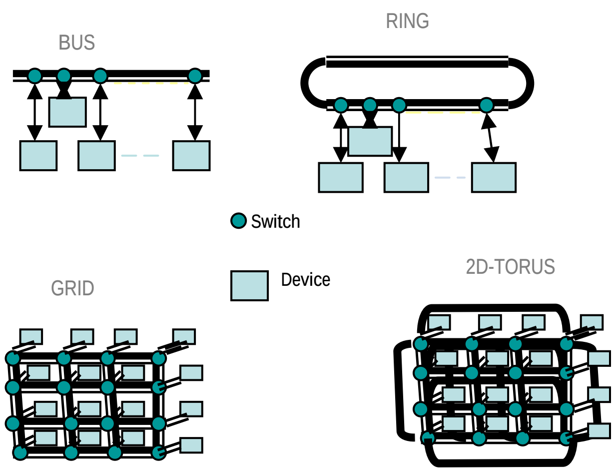

Рис.3. Топологии шины, кольца, грид и тора

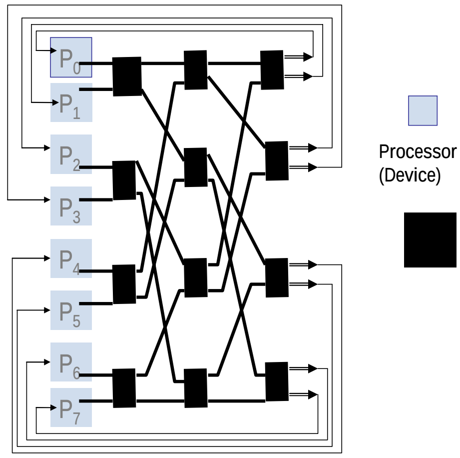

Рис. 4. Коммуникационная Ω-сеть

Пересекающиеся (“хордальные”) кольца могут иметь экономную отказоустойчивость (т.е. на один потенциальный отказ приходится один избыточный ресурс), но наращиваемость их затруднительна.

Полносвязный граф, рис. 5, 6, обладая многими привлекательными свойствами, совершенно не приемлем по стоимостным соображениям.

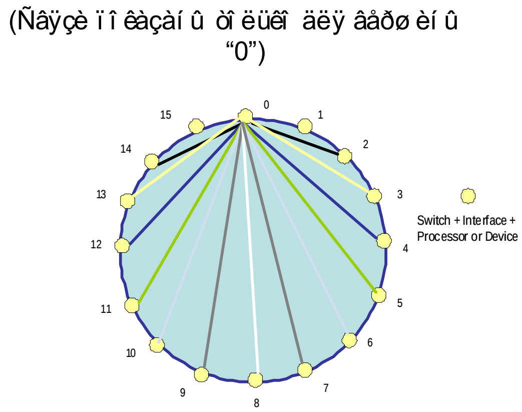

Рис. 5. Фрагмент полносвязного графа

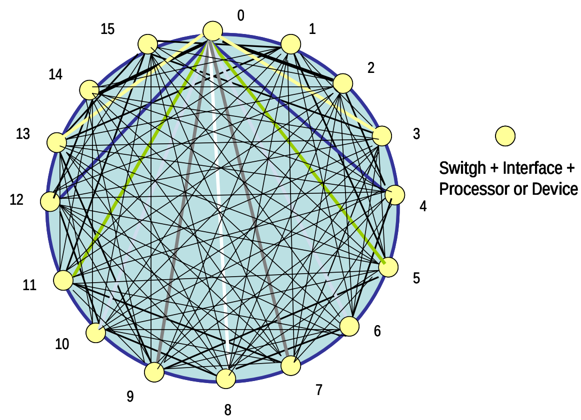

Рис.6. Полносвязный граф

Заключение

Рассмотренные здесь графы сохраняют в пределах каждой группы, определяемой табл. 1, постоянную степень вершин s, т.е. число портов в узлах прикладной системы. Это замечательное свойство открывает возможности масштабирования и введения отказоустойчивости в пределах группы, которую можно назвать однородным кластером. Решение поставленной в докладе задачи для бо́льших значений степени s расширит номенклатуру кластеров и число абонентов в них, предоставляя новые возможности для практических приложений. Привлекательность новой топологии вытекает не только из-за её преимуществ по числу связей и энергопотреблению сравнительно с полносвязными системами, но и из-за возможности снижения почти в $\sqrt{n}$ раз рабочей частоты при той же производительности, что и высокочастотные шинные системы. Основной "платой" за эти преимущества стало введение малоразмерных локальных коммутаторов-маршрутизаторов. Представляется, что в комплексном соотношении производительность-связность-наращиваемость-отказоустойчивость-энергопотребление новая топология достигает теоретического предела.

На рис. 8 - 10 дано качественное сравнение сетевых (кластерных) характеристик рассмотренных топологий.

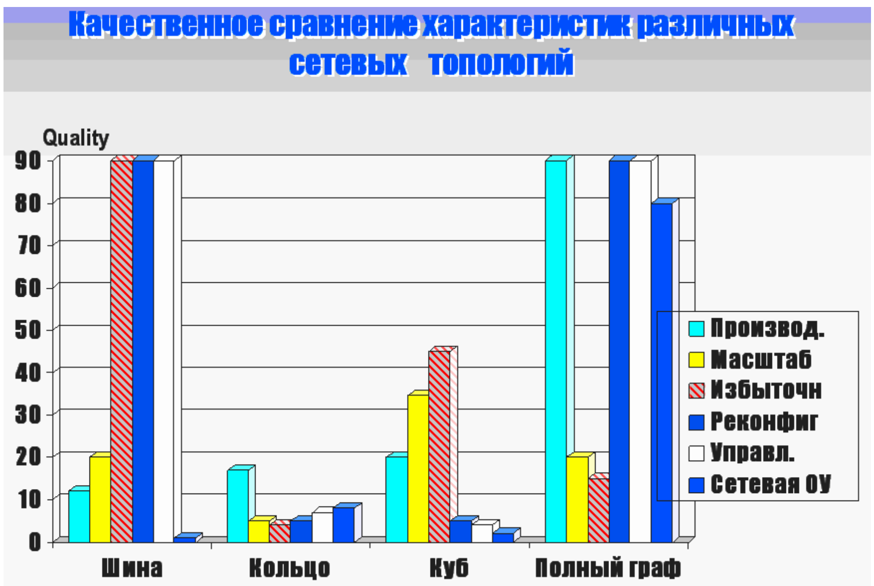

Рис. 7. Качественное сравнение сетевых характеристик различных топологий

Рис. 8. Качественное сравнение полного графа с кросс-кольцами

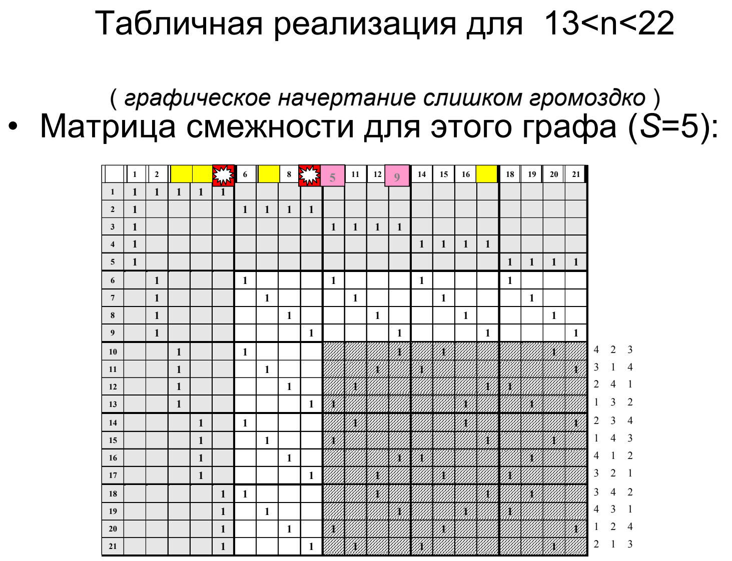

Рис. 9. Матрица инцидентностей для блок-схемы 21x21

Рис. 10. Сравнение качественных характеристик квазиполного и полного графов

Литература

1. Каравай М.Ф. Минимизированное вложение произвольных гамильтоновых графов в отказоустойчивый граф и реконфигурации при отказах. Часть II. Решетки и k-отказоустойчивость. // Автоматика и телемеханика. – 2005. – № 2. – С. 175-189.

2. Холл М. Комбинаторика. – М.: Мир, 1970. – 424 с.

3. Rosello J. SpaceWire Web Page, European Space Agency [Электрон. ресурс]. – Режим доступа: http://www.estec.esa.nl/tech/spacewire/.

4. Каравай М.Ф., Пархоменко П.П. , Подлазов В.С. Комбинаторные методы построения двудольных однородных неизбыточных квазиполносвязных графов (симметричных блок-схем). // Автоматика и телемеханика. – 2008. – №3 – С. 141-146.

Поступила в редакцию 5.03.2008

Рецензент: д-р техн. наук, проф. В.С. Харченко, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.