

УДК 021.8 + 025.1
ББК 78.34

О ТОПОЛОГИЧЕСКОЙ ОТКАЗОУСТОЙЧИВОСТИ МАСШТАБИРУЕМЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ¹

Мелентьев В. А.²

(ФГБУН Институт физики полупроводников
им. А.В. Ржанова СО РАН, Новосибирск)

Рассматриваются проблемы анализа топологической отказоустойчивости масштабируемой вычислительной системы (ВС) и обеспечения ее устойчивости к отказам заданной кратности. Предложен критерий топологической отказоустойчивости, напрямую связывающий топологию с потенциальным параллелизмом системы при заданной кратности допускаемых отказов. Определена взаимосвязь функций топологической масштабируемости и топологической отказоустойчивости систем. Показана обусловленность минимума топологической отказоустойчивости обхватом графа вычислительной системы. Модель параллельных вычислений, а также функции топологической отказоустойчивости и масштабируемости адаптированы к наличию уникальных узлов в информационной топологии решаемой задачи. Предложен способ конфигурирования отказоустойчивых подсистем при дефицитной топологической отказоустойчивости ВС, при этом обеспечение заданной для приложения кратности отказов достигается дублированием подсистем, сконфигурированных для меньшей, чем заданная, кратности отказов.

Ключевые слова: масштабируемые вычислительные системы, их топологическая отказоустойчивость.

¹ Работа выполнена при поддержке Российского фонда фундаментальных исследований, проект №14-07-00169а.

² Виктор Александрович Мелентьев, кандидат технических наук (melva@isp.nsc.ru).

1. Введение

Масштабирование системы, как правило, имеет целью увеличение производительности системы на задачах, критичных к времени их решения. Однако сопутствующий горизонтальному масштабированию экстенсивный рост количества используемых компонентов ведет к уменьшению времени наработки на отказ: порядка 20% вычислительной мощности высокопроизводительных систем теряется по причине возникших неисправностей и восстанавливается после них [2]. Поэтому в [15] особо выделена необходимость оперативного реконfigurирования мультипроцессорных сред и изоляции отказавших в процессе счета элементов, сохраняющего при этом эффект масштабирования; там же указывается на определяющее при этом значение системной топологии. Взаимосвязь и топологическая обусловленность масштабирования и отказоустойчивого функционирования ВС определяют потребность в совместном исследовании этих проблем именно в топологическом аспекте.

Несмотря на то, что известные проекты суперкомпьютеров класса «экзаскейл» (например, проект КНР [2]) предполагают использование межблочных шин с производительностью более 200 Гбит/с и латентностью менее 1,5 мкс, ключевой и наиболее насущной проблемой эксплуатации таких систем считается масштабирование и портирование ранее созданных приложений [3] с тем, чтобы они могли демонстрировать параллельную эффективность на вычислительном поле, не меньшем 20% используемых мощностей. В 2011 году это соответствовало 20 тысячам ядер, а к 2013 году это число увеличилось уже на порядок – до 200 тысяч [14].

Понятно, что в связи с нарастающей глобализацией информационных и вычислительных ресурсов интенсивность и кратность возникающих в таких системах отказов существенно возрастает. Эффективное их парирование требует адекватной и своевременной реакции на возникающие в связи с этим ситуации и заключается в поддержании функциональной и топологической целостности системы. Для исследования и реализации топологической целостности параллельных систем используют, как правило, две основные модели: модель с полной

k -отказоустойчивостью подсистем [20] и модель k -отказоустойчивых подсистем с амортизацией отказов.

Для первой модели характерна неизменность ранга p решаемой в подсистеме задачи даже при наличии k отказов, и k -отказоустойчивость достигается выявлением в графе ВС подграфа с минимизированной для заданных p и k вершинной избыточностью, достаточной для изоморфного (даже при удалении любых k вершин) вложения в него информационного графа задачи ранга p . Использование основанных на этой модели подходов в усовершенствовании и сопоставлении топологий вычислительных систем с позиции отказоустойчивости достаточно широко представлено как зарубежными [21, 22, 26], так и отечественными исследованиями [1, 4, 5]. В качестве основного критерия при этом используют, как правило, максимальный при заданной кратности отказов размер подсистемы, образуемой для решения задач с некоторой вполне определенной информационной топологией, или максимальную при заданном размере подсистемы кратность допускаемых отказов.

В данной работе проблема топологической отказоустойчивости вычислительных систем рассматривается в контексте модели с амортизацией отказов, согласно которой при возникновении в процессе решения задачи отказов ее ранг (размер подсистемы) может изменяться до приемлемого значения. При этом, в отличие от обговоренной выше модели, анализ обусловленной топологией отказоустойчивости ВС обычно осуществляют для системной сети связи в целом, не ограничивая ее рамками подсистем, образуемых для решения определенных, обладающих характерными информационными топологиями, задач и используя, как правило, сетевые показатели, статистически усредненные на множестве возможных при заданной кратности отказов конфигураций. Из критериев, имеющих отношение к проблеме отказоустойчивости, используют также критерий синхронизируемости [19], связывающий топологию сети с возможностью бесконфликтного (в пределах заданных временных ограничений, ограничений на приращение трафика и т.п.) общения произвольных вершин. В [18] предложен основанный на спектральных методиках подход к определению устойчивости этого критерия.

Отметим, что в силу экспоненциальной сложности точного вычисления большинство таких показателей являются стохастическими. Использование многокритериальной оптимизации в практике построения ВС, как правило, еще более усложняет процессы анализа и синтеза обладающих совокупностью заданных свойств топологий при том, что условия совместности/несовместности критериев также не безусловны и обосновываются стохастически. К тому же использование таких показателей в анализе топологической отказоустойчивости крупномасштабных систем представляется нам не вполне правомерным не только из-за их «комбинаторной взрывоопасности», связанной с непропорциональным (масштабированию системы и увеличению кратности отказов) ростом числа возможных конфигураций.

Во-первых, это связано с тем, что потенциалы эффективно масштабирования решаемых в системе задач совпадают с потенциалом системы лишь до некоторых пределов ее масштабирования. Поэтому, если для задачи существует предел эффективного распараллеливания, масштабирование системы сверх этого предела не приводит к аналогичному масштабированию ранга этой задачи, и число задач из решаемого набора, полностью использующих вычислительные (соответственно, и сетевые) ресурсы, с масштабированием системы будет сокращаться. Понятно при этом, что чем меньше в сравнении с системой размер подсистемы, тем в меньшей степени будет справедливым отнесение к ней сетевых показателей системы в целом. О недостаточной достоверности характеристики отказоустойчивости сетей статистически усредненными сетевыми критериями (в частности, средним расстоянием между узлами) говорится, например, в [17]. То же самое можно сказать и об использовании в сопоставлении топологий ВС таких критериев устойчивости к отказам как стабильность нагрузки узлов [25], стабильность диаметра, стабильность средних межузловых расстояний [17], стабильность связности [16] и о других традиционно используемых в анализе сетевой отказоустойчивости [13, 23] показателях.

Во-вторых, – и это, пожалуй, наиболее существенно – оценки влияния отказов на изменения потенциала параллелизма

вычислительной системы являются опосредованными через изменения ее сетевых характеристик. Отсутствие при этом формального соответствия значений сетевых показателей ВС ее потенциалу допускает, таким образом, не более чем качественную оценку топологической отказоустойчивости – на уровне «лучше/хуже».

Основная цель настоящей работы состоит в разработке критериев, не опосредованно, а напрямую связывающих топологию системы с потенциальным ее параллелизмом при заданной кратности допускаемых отказов. Приведено описание используемой при этом модели параллельных вычислений, обуславливающей получение требуемых при реализации приложений значений ускорения и эффективности в решении задачи ее рангом и предельным расстоянием между информационно смежными ветвями. Представлены способы описания лимитированных достижимостью и заданными значениями кратности отказов топологий ВС. Дано определение свойства топологической адекватности системы решаемым на ней задачам, предложен критерий оценки топологической отказоустойчивости, определена обусловленность топологической отказоустойчивости ВС обхватом ее графа. Установлена формальная взаимосвязь функций топологической отказоустойчивости и масштабируемости. Модель параллельных вычислений и эти функции адаптированы к наличию в приложениях уникальных ветвей. Предложен способ отказоустойчивой реализации приложений при дефицитной топологической отказоустойчивости системы.

2. Моделирование топологически отказоустойчивой масштабируемой вычислительной системы

2.1. ОПИСАНИЕ МОДЕЛИ

Зависимость масштабируемости параллельных систем и решаемых на них задач от топологии оценивается в [9] на модели параллельных вычислений, разделенной на две составляющие: первая (пп. 1–4) отнесена к параллельным приложениям и приписывает им свойства неограниченной распараллеливаемости, вторая (пп. 5–10) характеризует систему, ограничения параллелизма в которой обусловлены дефицитным быстродей-

ствием интерконнекта. Суть применения такой модели состоит не в том чтобы получить основанные на игнорировании архитектуры реальные оценки максимально возможного ускорения практически больших задач, а в формальном обосновании вносимых архитектурными компонентами системы (в данном случае – топологией интерконнекта) ограничений основного ее качества – параллелизма и в получении инструмента сопоставления топологий, исходя именно из этого качества. К сожалению, традиционно используемые в исследовании топологий интерконнекта сетевые критерии с указанным качеством систем непосредственного формального соответствия не имеют.

Понятно, что необходимость единообразного для всех топологий сопоставления по их влиянию на ограничения параллелизма потребовало выбора единой «точки отсчета» – эталонного приложения, обладающего идеальным в отношении параллелизма потенциалом. Только при этом условии допустимо считать, что все ограничения параллелизма в системе связаны исключительно с ее топологией, и соответственно, характеризовать топологию достигаемым с ее использованием потенциалом. Идеализация только приложений (пп. 1–4) обеспечивает возможность **единообразного** сопоставления топологий в отношении максимально достигаемого при их использовании потенциала параллелизма.

Заметим также, что принцип **единообразного** сопоставления систем по оценкам их пиковой производительности реализован выбором соответствующих тестов (например, LINPACK, HPL). Неполные соответствия тестов практическим задачам в части используемых параллельных алгоритмов, в части форматов представления данных и/или их объемов и т.п. не позволяют перенести полученные с их помощью абсолютные значения производительности систем на решаемые задачи. Однако эти данные повсеместно используют для **единообразной** рейтинговой оценки производительности систем в целом (в действующих их конфигурациях, включающих типы используемых процессоров, их число, интерконнектные топологии и технологии, интерфейсы программирования и т.п.).

Итак, исходные пункты описания модели содержат следующие обозначения: W и w – измеряемые временем объемы

вычислений при решении произвольной задачи на одном и на p процессорах вычислительной системы; Q и q – измеряемые информационными единицами (байтами) объемы подлежащих обмену данных, соответствующие одному и p задействованным в системе процессорам. Итак,

1. Задача допускает разбиение на произвольное число p информационно связанных параллельных ветвей: $1 \leq p \leq \infty$. Информационный граф распараллеленной на p ветвей задачи может быть нерегулярным, но обязательно связан.

2. Масштабирование данных в задаче с коэффициентом m увеличивает объем вычислений W и объем Q подлежащих обмену данных в m раз.

3. Общий объем вычислений W и объем Q подлежащих обмену данных при разбиении задачи на p параллельных ветвей не зависят от числа процессоров p и распределяются по ним равномерно: $w = W/p$ и $q = Q/p$.

4. Параллельный алгоритм не содержит скалярных фрагментов – это свойство вытекает из сформулированных выше пп. 1 и 3.

5. Все процессоры системы идентичны, их общее число n достаточно для реализации на них p параллельных ветвей, и первоначальное предварительное распределение входных данных по задействованным в параллельном приложении процессорам не требуется.

6. Топология ВС является регулярной¹ и неполносвязной.

7. Общие объемы W и Q не зависят от топологии сети связи и от используемой NT , и ограничения на минимальные объемы w и q отсутствуют.

8. Вычислительные и коммуникационные элементы ВС допускают совмещенную во времени работу.

9. Временные затраты на обмены пропорциональны расстояниям между информационно-связанными задачей вершинами графа ВС.

¹ Степени всех вершин графа ВС одинаковы.

10. Совокупность используемых в вычислительной системе топологии и NT гарантирует отсутствие сетевых коллизий и связанных с ними задержек.

Использование предложенной в [9] модели позволило установить формальную связь объемов W вычислений и Q обмениваемых данных распараллеливаемой на p ветвей задачи с предельно допускаемыми расстояниями δ между информационно смежными ее ветвями при заданных значениях ускорения – $S = T_1/T_p$ (здесь T_1 и T_p – времена решения задачи на одном и на p процессорах соответственно) или эффективности использования задействованных в решении процессоров – $E = S/p$:

$$(1a) \quad \delta_S(p) = \left\lfloor \frac{W}{S \cdot t_{NT}(Q/p)} \right\rfloor,$$

$$(1б) \quad \delta_E(p) = \left\lfloor \frac{W}{p \cdot E \cdot t_{NT}(Q/p)} \right\rfloor.$$

Естественно, что предельные (для рассматриваемых в конкретных ВС приложений) значения достижимостей δ различаются между собой тем значительнее, чем больше отличия в быстродействии используемых в этих ВС сетевых технологий (далее NT – Network Technology). Формулы (1a) и (1б), в которых t_{NT} – известная для NT зависимость времени задержки от объема передаваемой информации, достаточно наглядно демонстрируют это. Они же устанавливают зависимость предельного расстояния между информационно смежными ветвями распараллеливаемой задачи от заданного ускорения и присущих ей объемов вычислительных и обменных операций. Таким образом, реализуемые в исследуемой ВС¹ параллельные приложения могут быть классифицированы по значениям предельных для информационно смежных ветвей расстояний, соответствующих требуемым значениям ускорения S и/или эффективности E при условии топологической адекватности системы этим приложениям.

¹ Здесь и далее понятие ВС предполагает использование в ней вполне определенной NT с характерной для нее функцией $t_{NT}(Q/p)$.

Определение. Вычислительная система топологически адекватна задаче с присущими ей объемами W вычислительных операций и Q обрабатываемых данных, если топология системы позволяет сконфигурировать подсистему, в которой число процессоров p и предельное расстояние δ между информационно смежными процессорами соответствуют заданным значениям ускорения S и/или эффективности E ее решения.

Понятие топологической адекватности ВС и реализуемых на ней параллельных приложений неоднократно используется в последующем тексте и базируется на описанной выше модели и формальном ее выражении зависимостями (1а), (1б). Напомним, что если решение (W, Q) -задачи обусловлено получаемым для нее ускорением S , то число параллельных ветвей p в подсистеме определяет нижнюю границу числа используемых при этом процессоров, если же критерием является эффективность E использования процессоров, то число параллельных ветвей p в подсистеме определяет верхнюю границу их числа [9].

При необходимости одновременного удовлетворения критериям ускорения и эффективности, требования к топологии и к используемой системой сетевой технологии существенно ужесточаются: топология системы с действующей в ней NT должна быть привязана к информационной топологии задачи так, чтобы обеспечить превышение верхней (определяемой требуемой эффективностью E) границы числа p над его нижней (обусловленной заданным ускорением S) границей. Такое превышение может быть обеспечено, например, использованием более быстрой действующей NT , что в соответствии с (1а) позволяет использовать меньшее число процессоров. При конструктивной невозможности такого рода модификаций технологии сети связи и ее топологии, или при недостаточности этого для совместности E и S следует прибегнуть к поиску компромисса между ними, возможно, с масштабированием обрабатываемых задач и передаваемых между ее ветвями данных [9].

Как видно из (1а), (1б), связанное с отказами уменьшение числа исправных процессоров в реализующих параллельные приложения подсистемах отрицательно сказывается на ускорении, но не на эффективности решения параллельных задач. Поэтому в нашей работе по умолчанию предполагается, что

предельно допустимое для информационно смежных процессоров подсистемы расстояние δ определено из (1а), т.е. в соответствии с требуемым при отказоустойчивом решении задачи ускорением S .

Понятие изоморфизма в теории графов отражает взаимную однозначность (биективность) отображения вершин сопоставляемых графов одного порядка – при этом каждой вершине одного из них соответствует ровно одна вершина другого, и отношения смежности вершин и их образов совпадают. В теории вычислительных систем проблема изоморфизма состоит в выявлении в графе G системы подграфа, изоморфного информационному графу параллельной задачи $W(p)$ с заданным числом p параллельных ветвей. Имеющая при этом место дискретность в отношении изоморфизма вложения графа $W(p)$ в G (да/нет) не позволяет численно оценивать степень топологической адекватности вычислительной системы и реализуемых ею параллельных приложений. Такую возможность предоставляет введенный в [11] показатель масштабируемости $\mu(W, G_\delta)$ системы с G -топологией¹ в реализации на ней задачи с W -топологией при определяемом из (1а), (1б) значении достижимости δ :

$$\mu_\delta(W, G) = n_\delta(W, G)/n(G);$$

здесь $n(W, G_\delta) \equiv n_\delta(W, G)$ – порядок максимального, изоморфно вкладываемого в граф δ -достижимости системы G_δ , подграфа задачи с W -топологией.

Равенство единице показателя $\mu_\delta(W, G)$, говорит о том, что такой максимально вкладываемый подграф или является суграфом² графа G_δ , или изоморфен ему. Понятно, что если G_δ -граф полон, то масштабируемость в нем δ -задач с любой топологией максимальна и равна единице, т.е. он адекватен любым δ -задачам независимо от их информационной топологии. Меньшие значения этого показателя для рассматриваемых в системе задач соответствуют меньшей топологической ее адекватности этим задачам.

¹ Здесь топология системы задана графом G .

² Суграф, или остовный (вершинно-порожденный) подграф – часть графа, имеющая то же множество вершин, что и сам граф.

Ниже даны определение, способ получения и примеры использования графа δ -достижимости в исследовании топологической адекватности систем и решаемых на них задач.

2.2. ГРАФ δ -ДОСТИЖИМОСТИ

Напомним определение введенного в [6] понятия δ -достижимости как отношения смежности вершин в графе G_δ , расстояние (длина простого пути) между которыми в графе G вычислительной системы не превышает значения δ . Проиллюстрируем это простым примером сопоставления кольцевого графа G (рис. 1) и соответствующих ему графов 2- и 3-достижимости (рисунки 2 а и 2 б соответственно).

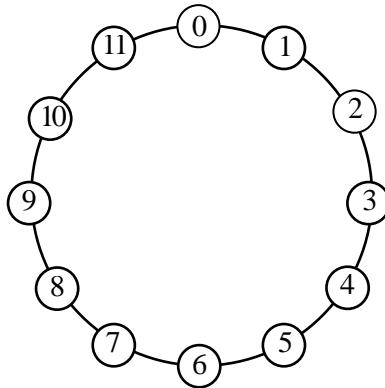


Рис. 1. Кольцевой граф G порядка $n(G) = 12$

Отметим, что все вершины рассматриваемого кольцевого графа дистанционно эквивалентны: i -окрестности любой пары вершин $u, v \in V$, вплоть до равного диаметру $d(G)$ индекса i , – попарно равномогны. Поэтому при построении графа δ -достижимости $G_\delta(G)$ можно ограничиться рассмотрением лишь одной вершины и ее расстояний до остальных вершин графа G . Напомним, что i -окрестность вершины $u \in V$ есть подграф $[u]_i \equiv G_i(u)$, индуцированный множеством вершин,

находящихся от вершины $u \in G$ на расстоянии¹ i . Множество вершин, составляющих i -окрестность вершины u называют ее i -окружением [8].

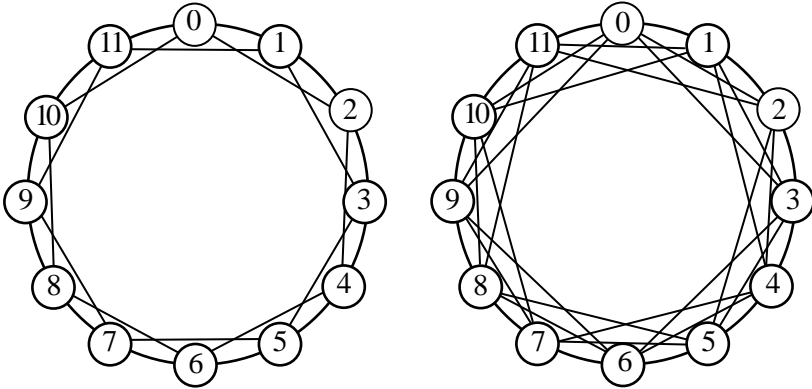


Рис. 2. Графы δ -достижимости кольцевого графа G :
 а – граф 2-достижимости G_2 ; б – граф 3-достижимости G_3

Построенная из любой вершины x кольца (рис. 1) минимально полная² его проекция $P_6(x)$ является 6-уровневой, и диаметр такого графа $d(G) = 6$. Для определенности в проекции (2) в качестве ракурсной выбрана вершина 0, она же составляет нулевой уровень этой проекции. Плотность рассматриваемого графа $\varphi(G) = \varphi(G_1) = 2$, это означает, что число p процессоров, которые могут быть задействованы в решении информационно полносвязных задач, лимитированных достижимостью $\delta = 1$, не превышает двух: $p \leq 2$. Для информационно неполносвязных задач это число может быть большим двух, вплоть до $p = 12$ – для задач с кольцевой топологией.

¹ Расстояние $d(u, v)$ между вершинами u и v равно длине кратчайшего соединяющего их пути.

² Проекция графа является полной, если она содержит в себе описания всех его вершин и инцидентных им ребер. Проекция с минимальным для ее полноты числом уровней является минимально полной.

графа G_2 порождается любыми тремя подряд расположенными вершинами $\{x, x + 1, x + 2\}$). Такое увеличение плотности означает, что на задачах с лимитируемой достижимостью $\delta = 2$ число параллельно используемых процессоров p может быть увеличено, как минимум, от трех (для полносвязных задач) до $n(G) = 12$ (для кольцевых задач).

Построенная из произвольной ракурсной вершины $x \in V$ проекция 3-окрестности «кольца» G имеет вид:

$$(5) P_1[x]_3 = x^{(x+3, x-3)}$$

Сжав (2) с учетом (3), (4) и (5), получим полную проекцию $P_{(0)}(G_3)$ графа 3-достижимости:

$$P_{(0)}(G_3) = 0^{(1(2^{(3,4,5,11)}, 3^{(2,4,5,6)}, 4^{(2,3,5,6,7)}, 10^{(7,8,9,11)}, 11^{(2,8,9,10)}), 2^{(1(3,4,10,11)}, 3^{(1,4,5,6)}, 4^{(1,3,5,6,7)}, 5^{(3,4,6,7,8)}, 11^{(2,8,9,10)}), 0^{(3^{(2,4,10,11)}, 2^{(1,4,5,11)}, 4^{(1,2,5,6,7)}, 5^{(2,4,6,7,8)}, 6^{(4,5,7,8,9)}, 9^{(6^{(3,4,5,7,8)}, 7^{(4,5,6,8,10)}, 8^{(5,6,7,10,11)}, 10^{(1,7,8,11)}, 11^{(1,2,8,10)}), 0^{(10^{(1(2,3,4,11)}, 7^{(4,5,6,8,9)}, 8^{(5,6,7,9,11)}, 9^{(6,7,8,11)}, 11^{(1,2,8,9)}), 11^{(1(2,3,4,10)}, 2^{(1,3,4,5)}, 8^{(5,6,7,9,10)}, 9^{(6,7,8,10)}, 10^{(1,7,8,9)})}$$

В этой проекции все вершины графа 3-достижимости определены двумя уровнями, и, хотя диаметр $d(G_3) = 2$, полнота описания ребер в G_3 и, соответственно, полнота его проекции, достигается не на двух, а, как и в графе 2-достижимости, на трех уровнях. Степень этого графа в сравнении с G_2 возрастает с четырех до шести, δ -плотность при этом повышается с $\varphi(G_2) = 3$ до $\varphi(G_3) = 4$.

Общеизвестно, что кольцевая топология обладает крайне ограниченными возможностями в реализации параллельных задач с отличной от кольцевой информационной топологией. Это подтверждается и рассмотренными примерами, из которых видно, что смещение вверх предельного расстояния δ между информационно смежными в задаче процессорами (от $\delta = 2$ до $\delta = 3$), которое при тех же (W, Q) -параметрах задачи может быть достигнуто только повышением быстродействия используемой в системе сетевой технологии, но это даст лишь незначительное увеличение параллелизма приложений: для информационно полносвязных задач число параллельных ветвей может быть увеличено с трех (при $\delta = 2$) всего лишь до четырех ($\delta = 3$).

3. Топологическая отказоустойчивость масштабируемых вычислительных систем

Как мы уже указывали в описании используемой здесь модели (раздел 2.1), реализуемые исследуемой ВС параллельные приложения могут быть классифицированы по значениям предельных для информационно смежных ветвей расстояний δ . Подобные ограничения длины пути, связанные с недостаточным для актуальности информационных обменов быстродействием сетевых технологий, имеют место и в других, например, в распределенных информационных системах. Поэтому ценность исследований, отнесенных к сети в целом и игнорирующих при этом потребности актуализации в ней информационных обменов, с увеличением масштаба сетей связи существенно снижается. Традиционное представление сетей связи графами, в которых отношения смежности вершин соответствуют физической смежности сетевых узлов, также не способствует исследованию сетей в реализации лимитированных транзитных взаимодействий, доля которых в общем объеме информационных обменов растет непропорционально сетевым масштабам. В описанных в разделе 2.2 графах δ -достижимости смежность вершин обусловлена заданным предельно допускаемым расстоянием $\delta \geq 1$ между ними, и это позволяет коррелировать характеристики сети допускаемыми в ней задержками. Кроме того, применение таких графов в конфигурировании в ВС подсистем позволяет исключить возможность превышения допустимых задержек между информационно смежными ветвями задач.

Для исследования топологических аспектов устойчивости ВС к отказам кратности l в данной работе мы используем графы $\delta(k)$ -достижимости, в которых смежность вершин, в отличие от графов δ -достижимости, лимитирована не только допускаемым расстоянием δ между ними, но и числом $k = l + 1$ не превышающих этого расстояния независимых путей. Отметим при этом, что понятие отказоустойчивости имеет различную окраску при отнесении его к сети связи ВС или к реализующей параллельное приложение подсистеме. В первом случае вполне допустимо использование традиционных сетевых показателей, отнесенных,

однако, не только к исходному графу ВС (с $\partial = 1$), а к производным от него графам $\partial(k)$ -достижимости (с $\partial > 1, k \geq 1$). Если при этом в качестве критерия $\partial(l)$ -отказоустойчивости сети связи ВС выбрать, например, связность, то отношение порядка $\partial(k)$ -компоненты связности (компоненты связности графа $\partial(k)$ -достижимости) к порядку исходного графа ВС даст достаточно наглядное представление о присущем рассматриваемой топологии ВС уровне ее отказоустойчивости. Однако в данной работе нас интересуют не топологические возможности системы в целом, а возможности отказоустойчивой реализации в ней параллельных приложений, и именно в этом аспекте приведено изложение последующего материала.

3.1. ГРАФ $\partial(k)$ -ДОСТИЖИМОСТИ

Определим граф $\partial(k)$ -достижимости $G_{\partial(k)}(G)$, он же – граф $k(\partial)$ -соединимости $G_{k(\partial)}(G)$ [6], как надграф графа $G(V, E)$, дополненный ребрами между вершинами $u, v, \in V$ при наличии между ними в графе G не менее k независимых путей с длиной, не превышающей значение заданной достижимости ∂ . Очевидно, что число k таких путей (соединимость информационно смежных в подсистеме вершин подграфа системы) должно по меньшей мере превышать заданную кратность l отказов: $k \geq l + 1$. В этом случае при удалении l каких-либо вершин из соответствующего выполняемой задаче подграфа ВС любые две из оставшихся в нем информационно смежных вершин по-прежнему останутся соединенными путем с длиной, не превышающей предельной для этой задачи достижимости ∂ .

Построить полную проекцию графа $\partial(k)$ -достижимости можно сжатием всего одной полной проекции исходного графа. При этом критерием помещения вершины на j -й уровень выстраиваемой проекции будет не просто присутствие ее в составе i -окружений (при $1 \leq i \leq \partial$) вершин $(j - 1)$ -го уровня, а ее повторяемость с кратностью $m \geq k$. Вместе с этим должен быть проведен анализ на независимость всех m ∂ -путей между вершинами $(j - 1)$ -го уровня и порожденными ими вершинами j -го уровня сопоставлением промежуточных на каждом пути вершин, при этом пути с непустым множеством совпадений считаются зави-

симыми. Большее или равное k число выявленных таким образом независимых путей между вершинами u и v в графе G соответствует наличию ребра между ними в графе $G_{\hat{\alpha}(k)}(G)$.

Как мы уже упоминали выше, все вершины рассматриваемого нами кольцевого графа G – дистанционно эквивалентны, поэтому для построения графа $G_{\hat{\alpha}(k)}(G)$ достаточно единственной проекции графа G . Однако более наглядным будет построение графа $\hat{\alpha}(k)$ -достижимости $G_{\hat{\alpha}(k)}(G)$ с использованием $n(G)$ $\hat{\alpha}$ -уровневых проекций исходного графа $G(V, E)$. В этом случае единственный, стоящий над нулевым, уровень проекции $P_1[x]_i \mid i \in \infty^{\hat{\alpha}}$, определяет перечень смежных по отношению к x вершин в графе $G_{\hat{\alpha}(k)}(G)$. Здесь $\infty^{\hat{\alpha}}$ – последовательность всех натуральных чисел от единицы до $\hat{\alpha}$. Далее условие перебора i -окрестностей от $i = 1$ до $i = \hat{\alpha}$ в проекции $P_1[x]_i \mid i \in \infty^{\hat{\alpha}}$ для компактности опускаем – $P_1[x]_i$.

Из приведенной в (2) проекции $P_6(0)$ понятно, что граф $G_{\hat{\alpha}(2)}(G)$ $\hat{\alpha}(2)$ -достижимости исходного (рис. 1) кольцевого графа G при $\hat{\alpha} < 6$ не содержит ребер и, таким образом, является пустым – $\hat{\alpha} < 6 \Rightarrow \forall x \in V P_1[x]_i = x$. Единичное значение плотности при этом указывает на отсутствие в системе возможностей конфигурирования отказоустойчивых подсистем для задач с $\hat{\alpha} < 6$. Если же допускаемое задачей расстояние между информационно смежными процессорами $\hat{\alpha} = 6$, то $P_1[x]_i = x^{(x+6)}$, граф $G_{6(2)}(G)$ образован шестью 2-вершинными (со смежными в этом графе вершинами x и $x + 6$) компонентами связности – плотность графа и максимальная степень параллелизма полносвязных задач при этом равны двум. Для $\hat{\alpha} = 8$ получим $P_1[x]_i = x^{(x+6, x+7, x+8, x-7, x-8)}$. Полная проекция графа $8(2)$ -достижимости, построенная из нулевой вершины, при этом имеет вид:

$$P_{(0)}(G_{8(2)}) = 0^{(4^{(8(1,2,3)}_9(1,2,3,5)_{10}(2,3,5,6)_{11}(3,5,6,7)}_5(6,7,8,9)_{9(1,2,3,4)_{10}(2,3,4,6)_{11}(3,4,6,7)}),}$$

$$0^{(6^{(1(5,7,8,9)}_2(7,8,9,10)_{10}(2,3,4,5)_{11}(3,4,5,7)}_7(5,6,8,9)_{2(6,8,9,10)}_3(8,9,10,11)_{11}(3,4,5,6)}),}$$

$$0^{(8^{(1(5,6,7,9)}_2(6,7,9,10)_{3(7,9,10,11)}_4(9,10,11)}),}$$

Соответствующие полученным при $\hat{\alpha} = 6$ и $\hat{\alpha} = 8$ проекциям графы $G_{6(2)}(G)$ и $G_{8(2)}(G)$ приведены на рис. 3а и 3б. Как нетруд-

но заметить, диаметры их – $d(G_{6(2)}) = \infty$ и $d(G_{8(2)}) = 2$, а плотности – $\varphi(G_{6(2)}) = 2$ и $\varphi(G_{8(2)}) = 3$.

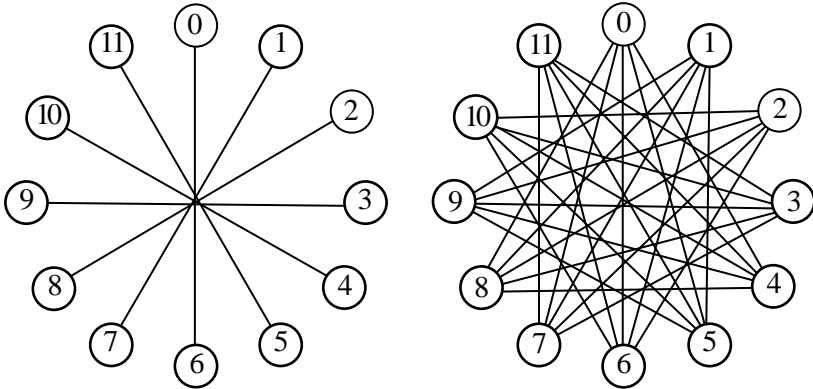


Рис. 3. Графы $\partial(2)$ -достижимости кольцевого графа G :
 а – граф $6(2)$ -достижимости $G_{6(2)}(G)$;
 б – граф $8(2)$ -достижимости $G_{8(2)}(G)$.

3.2. ОБХВАТ ГРАФА ВС И ЕЕ ТОПОЛОГИЧЕСКАЯ ОТКАЗОУСТОЙЧИВОСТЬ

Рассмотренные в предыдущем разделе примеры подтверждают общеизвестный факт крайне низкого потенциала кольцевых топологий в реализации на ВС отказоустойчивых приложений: даже для топологически адекватных системе задач возможность отказоустойчивой их реализации может появиться только при сравнимой с порядком «кольца» достижимости: $\partial \geq \lfloor n/2 \rfloor$. Это существенно ограничивает возможности масштабирования требующих отказоустойчивости приложений даже при использовании в ВС сверхбыстродействующих сетевых технологий: каким бы ни было быстродействие NT в кольцевой ВС, число независимых цепей между любыми двумя вершинами не превысит двух, таким образом, кратность допускаемых в ней отказов не может превышать единицы: $l \leq 1$.

Однако из этих же примеров видно, что ограничения топологической отказоустойчивости ВС связаны с обхватом¹ ее графа, который для кольцевых систем имеет предельное, равное порядку «кольца» значение. Но очевидно, что подобная обусловленность имеет место и для систем с любой другой топологией.

Утверждение. Обхват g графа G вычислительной системы не должен превышать удвоенного для отказоустойчивого решения задачи W значения достижимости $\delta(W)$:

$$(3) \quad g(G) \leq 2\delta.$$

Доказательство. Если обхват $g(G)$ графа ВС превышает удвоенное значение регламентируемой задачей достижимости δ , то в этом графе по определению обхвата не найдется и пары вершин $u, v \in V$, связанных хотя бы двумя независимыми цепями с длинами, не превышающими этой достижимости: $\delta < \lceil g(G)/2 \rceil \Rightarrow k(u, v) \leq 1$. Следовательно, граф $\delta(k)$ -достижимости $G_{\delta(k)}$ будет пустым при отличной от нуля кратности l отказов, порядок компоненты связности в нем равен единице, что отрицает возможность конфигурирования отказоустойчивых подсистем с $\delta < \lceil g(G)/2 \rceil$ в принципе.

Понятно, что для набора реализуемых в ВС отказоустойчивых приложений обхват графа ВС обусловлен минимальным в наборе таких приложений значением достижимости δ_{\min} :

$$g(G) \leq 2\delta_{\min}.$$

Из (3) и отсутствия кратных ребер между вершинами графа ВС ясно, что для задач, в решении которых из-за дефицитного быстрогодействия NT не может быть обеспечено большее единицы значение достижимости – $\delta > 1$, свойство топологической отказоустойчивости не реализуемо в принципе: даже при использовании полностью связной топологии: $\delta = 1 \Rightarrow l = 0$. Поэтому для того чтобы, например, гиперкубическая ВС была способна парировать однократные отказы, быстрогодействие ее NT должно быть достаточным для обеспечения достижимости $\delta \geq 2$.

¹ Обхватом графа называется длина его кратчайшего цикла.

Очевидно, что при кратности $l \geq 1$ отказов ($k \geq 2$) обусловленное задачами ограничение (3) обхвата графа ВС дополняется ограничением снизу его степени, которая, как и обхват, также влияет на диаметр графа и на его порядок. Это объясняется тем, что в s -регулярном графе с заданными значениями порядка n и степени s наибольшей компактностью¹ обладают графы с наибольшими обхватами [12], и тем, что число k независимых цепей между двумя вершинами регулярного графа не может превышать его степени: $k \leq s$. Поэтому утверждение о том, что «снижение диаметра графа межмашинных связей ведет к увеличению эффективности и повышению надежности системы...» [12], может быть справедливым лишь в отношении эффективности, но не в отношении отказоустойчивости, так как при заданных значениях порядка и степени графа стремление к минимизации его диаметра $d_{\min}(G)$ требует увеличения обхвата до максимального значения $g_{\max}(G)$, которое, как известно [7], равно удвоенному диаметру $g_{\max}(G) = 2d_{\min}(G)$. Это нарушает условие (3) для всех задач с меньшей этого диаметра достижимостью $\partial \leq d_{\min}(G)$ и, таким образом, приводит к полной утрате свойства топологической отказоустойчивости системы при решении таких задач.

3.3. ФУНКЦИЯ ТОПОЛОГИЧЕСКОЙ ОТКАЗОУСТОЙЧИВОСТИ ВС

В принципе, введенная в [9] функция топологической масштабируемости $\mu_k(G)$ способна дать представление и о топологической отказоустойчивости вычислительной системы. Для этого необходимо функцию плотности $\varphi(G_\partial)$ графа G системы заменить определенной в разделе 2.3 плотностью $\varphi(G_{\alpha(k)})$ и аргументировать $\varphi(G_{\alpha(k)})$ не числом $n \equiv n(G)$ узлов в ней (при нулевой кратности отказов $l = 0$), как это делается в $\mu_k(G)$, а кратностью $l = k - 1$ отказов. Изменив при этом обозначение

¹ Компактные структуры ВС определены как s -регулярные заданного порядка графы с минимально возможным диаметром.

функции топологической масштабируемости с $\mu_c(G)$ на $\mu_{c,l}(n)$, получим:

$$(6) \quad \mu_{c,l}(n) = \varphi(G_{\partial(k)})/n(G).$$

Давая, однако, количественную характеристику потенциала параллелизма ВС (какая часть общего числа $n(G)$ процессоров может быть использована в l -отказоустойчивом решении ∂ -задач) функция топологической масштабируемости отказоустойчивой ВС (6) дает лишь качественное представление об изменении этого потенциала в сравнении с изначальным (при $l=0$) потенциалом – $\varphi_c(G)$, который у топологически различающихся ВС одного порядка, как правило, тоже различен.

Поэтому топологическую l -отказоустойчивость $\theta_{n,\partial}(l)$ системы с G -топологией ($n \equiv n(G)$) в решении ∂ -задач определим отношением плотности $\varphi_{\partial(k)}(G) \equiv \varphi(G_{\partial(k)})$ графа $\partial(k)$ -достижимости $G_{\partial(k)}$, соответствующей определенной для задач из (1а), (1б) достижимости ∂ , и заданной кратности $l \geq 1$ отказов ($k = l + 1$), к исключаяющей наличие отказов ($l = 0$) плотности $\varphi_c(G) \equiv \varphi(G_\partial)$ графа ∂ -достижимости G_∂ :

$$(7) \quad \theta_{n,\partial}(l) = \varphi_{\partial(k)}(G)/\varphi_c(G).$$

Сравнивая (6) и (7), нетрудно заметить, что для заданных значений n , ∂ и l функция топологической l -отказоустойчивости $\theta_{n,\partial}(l)$ равна отношению функции топологической масштабируемости l -отказоустойчивой ВС (6) к функции топологической масштабируемости системы $\mu_c(G) = \varphi_c(G)/n(G)$, определяемой в предположении отсутствия ($l = 0$) отказов:

$$\theta_{n,\partial}(l) = \mu_{c,l}(n)/\mu_c(G).$$

В отличие от традиционно используемых в сопоставлении ВС с различными топологиями сетевых показателей (диаметра, ширины бисекции), позволяющих судить лишь о качестве (лучше/хуже, больше/меньше) влияния топологии на потенциал распараллеливаемости задач в системе, показатели топологической отказоустойчивости ВС – (6) и топологической ее масштабируемости – (7) позволяют численно оценивать нижнюю (при распараллеливании в системе информационно ∂ -полносвязных задач) границу такого потенциала при возникновении в процессе их решения отказов кратности l . Следует помнить, однако,

что такая оценка справедлива лишь при паритетности (идентичности) всех параллельных ветвей задачи.

3.4. НЕПАРИТЕТНОСТЬ ВЕТВЕЙ ЗАДАЧИ И ТОПОЛОГИЧЕСКАЯ ОТКАЗОУСТОЙЧИВОСТЬ ВС

В практике параллельных вычислений далеко не редкими являются задачи, в которых отдельные параллельные ветви функционально или топологически отличны от остальных и в этом смысле такие ветви уникальны. К примеру, в задачах со звездной топологией уникальной является центральная вершина «звезды». То же можно сказать и в отношении используемых в решении задач схем информационного взаимодействия: наличие центрального узла свойственно трансляционной и коллекторной схемам обменов. Понятно, что для наделения свойством l -отказоустойчивости сконфигурированных под такие задачи подсистем процессоры, соответствующие x уникальным ветвям, должны быть l -кратно продублированы. Поэтому в сравнении с плотностью $\varphi(G_{\alpha(k)})$ число процессоров, задействованных в решении таких задач при их информационно полной связности, уменьшится на произведение $x \cdot l$ – соответственно уменьшаются значения топологической l -отказоустойчивости и топологической масштабируемости l -отказоустойчивой системы:

$$\theta_{\alpha(l)}(G) = (\varphi_{\alpha(k)}(G) - l \cdot x) / \varphi_{\alpha}(G),$$

$$\mu_{\alpha(l)}(G) = (\varphi_{\alpha(k)}(G) - l \cdot x) / n(G).$$

Для информационно неполносвязных задач, параллельные алгоритмы решения которых предполагают наличие уникальных ветвей ($x > 0$), информационный граф W исследуемой задачи следует дополнить такими $l \cdot x$ вершинами и инцидентными им ребрами, что окружение дублирующей вершины совпадает с окружением дублируемой уникальной вершины. Модифицированный таким образом граф задачи W обозначим W_l . Для определения топологической отказоустойчивости и масштабируемости $\mu_{\alpha(k)}(W, G) \equiv \mu(W, G_{\alpha(k)})$ системы при l -отказоустойчивой реализации решения такой W -задачи следует определить порядок $n(W_l, G_{\alpha(k)}) \equiv n_{\alpha(k)}(W_l, G)$ максимального изоморфного графу W_l подграфа в графе $\partial(k)$ -достижимости $G_{\alpha(k)}$, здесь $-k = l + 1$. Способ выявления изоморфного вложения информационного

графа задачи в такие графы предложен и подробно рассмотрен в [6] и [8]. Тогда топологическая отказоустойчивость и масштабируемость системы в решении таких задач определяются выражениями:

$$(8) \quad \theta_{\alpha(l)}(W, G) = (n(W_l, G_{\alpha(k)}) - l \cdot x) / \varphi_{\partial}(G),$$

$$(9) \quad \mu_{\alpha(l)}(W, G) = (n(W_l, G_{\alpha(k)}) - l \cdot x) / n(G).$$

Из (8), (9) ясно, что использование алгоритмов с непаритетными ветвями, существенно ухудшает потенциал отказоустойчивого и параллельного решения задач даже при их полной топологической адекватности системе.

Резюмируя вышесказанное, отметим, что требование топологической отказоустойчивости в решении задач ограничивает потенциал параллелизма в их решении тем в большей степени, чем выше поднята планка кратности допускаемых при решении задач отказов, так как $n(W_{l+1}, G_{\alpha(l+2)}) < n(W_l, G_{\alpha(l+1)})$ (для полностью связанных задач – $\varphi(G_{\alpha(k+1)}) < \varphi(G_{\alpha(k)})$). При этом может случиться так, что необходимое для обеспечения ускорения S задачи число процессоров p , соответствующее достижимости ∂ и заданной кратности l отказов, топологически не может быть обеспечено: $p > n(W_l, G_{\alpha(l+1)})$. Один из возможных способов решения этой проблемы приведем ниже.

Пусть топология ВС задана графом G , информационная топология решаемой задачи – графом W . Кратность допускаемых при решении этой задачи отказов отлична от нуля. Необходимое для обеспечения топологической l -отказоустойчивости число k независимых путей между процессорами, реализующими информационно смежные в задаче ветви, $k = l + 1$. Пусть полученное из (1а) минимальное число p процессоров, необходимое для получения заданного в решении данной задачи ускорения превышает порядок максимального изоморфного графу W_l подграфа в графе $\partial(k)$ -достижимости $G_{\alpha(k)} - p > n(W_l, G_{\alpha(k)})$. Пусть при этом существует некоторая отличная от нуля величина $y \leq l$, дополняющая до l значение кратности $l - y$ отказов, при которой $p \leq n(W_{l-y}, G_{\alpha(k-y)})$. Тогда если в графе $G_{\alpha(k-y)}$ может быть организовано $y + 1$ подграфов W_{l-y} с $(l - y)$ -отказоустойчивой для рассматриваемой задачи топологией, то устойчивость к кратности l отказов в решении задачи может быть достигнута u -кратным

резервированием топологически $(l - y)$ -отказоустойчивых подсистем.

4. Заключение

Наращивание ресурсов в системе и стремление к увеличению их загрузки неизбежно приводят к увеличению вероятности отказов и их кратности [24]. В этой связи основное требование к масштабируемым системам, состоящее в том, что они должны допускать наращивание ресурсов, обеспечивающее пропорциональный прирост производительности без глобальной перестройки архитектуры, следует дополнить требованием повышения устойчивости к отказам или сохранения ее как минимум на уровне, предшествующем масштабированию.

Потенциал параллелизма системы во многом ограничен топологическим потенциалом обеспечения информационно-логической целостности решаемых задач. Традиционно используемые при сопоставлении вычислительных систем с различными топологиями сетевые характеристики квалифицируют лишь, собственно, топологии без учета потенциалов параллелизма, достигаемых системой при их использовании, и могут быть, таким образом, не более чем качественными (лучше/хуже). То же можно сказать и об использовании этих характеристик в сопоставлении отказоустойчивости топологически отличающихся систем.

В работе дано краткое описание модели, формально связывающей объемные параметры параллельной задачи (объемы вычислений и обрабатываемых данных) и требуемые значения критериев актуальности (ускорения) или эффективности ее решения, с необходимыми при этом и обусловленными быстрым действием используемой в системе сетевой технологии числом процессоров и предельными расстояниями (достижимостью) между ними. Основанное на такой модели свойство топологической отказоустойчивости ВС предложено оценивать относительно, при увеличении кратности l отказов, изменением плотности графа $\partial(k)$ -достижимости, где $k = l + 1$ – число независимых с длиной, не превышающей ∂ , цепей между процессо-

рами, соответствующими информационно смежным ветвям задач.

Показана формальная обусловленность топологической отказоустойчивости системы обхватом ее графа. Установлена формальная взаимосвязь топологической отказоустойчивости и топологической масштабируемости систем. Модель параллельных вычислений, функции топологической отказоустойчивости и топологической масштабируемости адаптированы к задачам, в которых отдельные параллельные ветви функционально или топологически отличны от остальных и в этом смысле являются уникальными. Рассмотрен случай с дефицитной в сравнении с заданной топологической отказоустойчивостью ВС, и предложен способ решения этой проблемы, когда обеспечение заданной для приложения кратности отказов может быть достигнуто дублированием подсистем, сконфигурированных для меньшей кратности.

Результаты работы могут быть полезны при исследовании деградации параллелизма действующих в условиях отказов систем, при выборе или генерации топологий вычислительных систем, ориентированных на отказоустойчивую реализацию набора решаемых задач, а также при распараллеливании таких задач и при конфигурировании соответствующих им подсистем. Использование полученных в работе результатов открывает также возможности детерминированного исследования и проектирования крупномасштабных информационных сетей с заданной устойчивостью к отказам в отношении сохранности и актуальности предоставления тех или иных информационных ресурсов, а также оптимизированного в соответствии с этим размещения их в сети.

Литература

1. АБРОСИМОВ М.Б. *Графовые модели отказоустойчивости*: дисс. докт. физ.-мат. наук. – URL: https://cs.msu.ru/sites/cmc/files/theses/doktorskaya_abrosimov.pdf (дата обращения: 18.10.2016).
2. ГЕРГЕЛЬ В.П., ЛИНЁВ А.В. *Проблемы и перспективы достижения экзафлопного уровня производительности суперкомпьютерных систем* // Информационные технологии. Вестник Нижегородского университета им. Н.И. Лобачевского. – 2012. – №3(1). – С. 189–198.
3. *Китай планирует построить суперкомпьютер уровня «экзаскейла» в 2020 году*. – URL: http://www.thg.ru/technews/20160503_110013.html/esk_tex.pdf (дата обращения: 03.06.2016).
4. КАРАВАЙ М.Ф. Применение теории симметрии к анализу и синтезу отказоустойчивых систем // Автоматика и телемеханика. – 1996. – №6. – С. 159–173.
5. КАРАВАЙ М.Ф., ПОДЛАЗОВ В.С. *Расширенный обобщенный гиперкуб как отказоустойчивая системная сеть для многопроцессорных систем* // Управление большими системами. – 2013. – Вып. 45. – С. 344–371.
6. МЕЛЕНТЬЕВ В.А. *Вложение подсистем, лимитирующих длину и число путей между вершинами графа вычислительной системы* // Управление большими системами. – 2015. – Вып. 47. – С. 212–246.
7. МЕЛЕНТЬЕВ В.А. *Компактные структуры вычислительных систем и их синтез* // Управление большими системами. – 2011. – Вып. 32. – С. 241–261.
8. МЕЛЕНТЬЕВ В.А. *Предельное конфигурирование подсистем в гиперкубических вычислительных системах* // Информационные технологии и вычислительные системы. – М.: Федеральный исследовательский центр «Информатика и управление» РАН, 2015. – №2. – С. 20–30.
9. МЕЛЕНТЬЕВ В.А. *О топологической масштабируемости вычислительных систем* // Управление большими системами. – 2015. – Вып. 58. – С. 115–143.

10. МЕЛЕНТЬЕВ В.А. *Формальные основы скобочных образов в теории графов* // Труды Второй Междунар. конф. «Параллельные вычисления и задачи управления» (РАСО), 4-6 октября 2004 г., Москва, Россия. – 2004. – С. 694–706.
11. МЕЛЕНТЬЕВ В.А., ШУБИН В.И., ЗАДОРОЖНЫЙ А.Ф. *Топологическая масштабируемость гиперкубических параллельных систем и задач* // ISJ Theoretical & Applied Science. – 2015. – №11. – С. 122–129.
12. ПИНЧУК В.П. *Базовые графы для построения топологии управляющих сетей многопроцессорных систем* // Искусственный интеллект. – 2004. – №4. – С. 46–58.
13. СНАРСКИЙ А.А., ЛАНДЭ Д.В. *Моделирование сложных сетей. Учебное пособие.* – Киев: НТУУ «КПИ», 2015. – 212 с. – URL: <http://freescb.info/sites/freescb.info/files/mss-new.pdf> (дата обращения: 20.10.2016)
14. СТЕГАЙЛОВ В.В., НОРМАН Г.Э. *Проблемы развития суперкомпьютерной отрасли в России: взгляд пользователя высокопроизводительных систем* // Программные системы: теория и приложения. – 2014. – №1(19). – С. 111–152.
15. *Эксафлопные технологии. Концепция по развитию технологии высокопроизводительных вычислений на базе суперэвм эксафлопного класса (2012–2020 гг.).* – URL: http://filearchive.cnews.ru/doc/2012/03/esk_tex.pdf (дата обращения: 01.06.2016).
16. ADHIKARI N., TRIPATHY C.R. *The Folded Crossed Cube: A New Interconnection Network for Parallel Systems* // Int. Journal of Computer Applications. – July, 2010. – Vol. 4(3). – P. 43–50. – URL: <http://www.ijcaonline.org/volume4/number3/rxc3871147.pdf> (дата обращения: 20.10.2016)
17. ALBERT R., JEONG H., BARABASI A. *Error and attack tolerance of complex networks* // Nature. – 2000. – Vol. 406. – P. 378–382. – URL: <http://barabasi.com/f/77.pdf> (дата обращения: 20.10.2016).
18. BARAHONA M., PECORA L.M. *Synchronization in Small-world Systems* // Phys. Rev. Lett. – 2002. – Vol. 89, Iss. 5. – URL: <https://doi.org/10.1103/PhysRevLett.89.054101> (дата обращения: 20.10.2016).

19. DONETTI L., HURTADO P.I., MUÑOZ M.A. *Entangled networks, synchronization and optimal network topology* // Phys. Rev. Lett. – 2005. – Vol. 95, Iss. 18. – URL: <https://doi.org/10.1103/PhysRevLett.95.188701> (дата обращения: 20.10.2016).
20. HAYES J.P. *A graph model for fault-tolerant computing system* // IEEE Trans. Comput. – 1976. – Vol. C-25, No. 9. – P. 875–884.
21. HARARY F., HAYES J.P. *Edge fault tolerance in graphs* // Networks. – 1993. – No. 23. – P. 135–142.
22. HO T.-Y., SUNG T.-Y., HSU L.-H. *A note on edge fault tolerance with respect to hypercubes* // Applied Mathematics Letters. – 2005. – No. 18. – P. 1125–1128.
23. KLAU G.W., WEISKIRCHER R. *Robustness and Resilience* // Network analysis. – 2005. – P. 417–437. – URL: <http://homepages.cwi.nl/~klaus/pubs/kw-rr:2005.pdf> (дата обращения: 20.10.2016).
24. MELENT'EV V.A. *Reliability of elements of the computing system and its fault tolerance* // ISJ Theoretical & Applied Science. – 2014. – №9(17). – P. 34–45.
25. NEWMAN M.E.J. *A measure of betweenness centrality based on random walks* // Social Networks. – 2005. – Vol. 27. – P. 39–54. – URL: <https://pdfs.semanticscholar.org/0a57/5498f9e6bc0cc43b977c6e952101f89be90c.pdf> (дата обращения: 20.10.2016).
26. ZIMMERMAN G.W., ESFAHANIAN A.-H. *Chordal rings as fault-tolerant loops* // Discrete Applied Mathematics. – 1992. – No. 37/38. – P. 563–573.

ON TOPOLOGICAL FAULT-TOLERANCE OF SCALABLE COMPUTING SYSTEMS

Victor Melentiev, Rzhanov Institute of Semiconductor Physics Siberian Branch of RAS, Novosibirsk, Cand. Sc., senior research associate (melva@isp.nsc.ru).

Abstract: Problems of the analysis of topological fault tolerance of the scalable computing system and ensuring its sustainability to fault of the given multiplicity are considered. The measure of topological fault tolerance is offered, which connects the computing system topology with its potential parallelism for the given fault multiplicity. The relationship between the functions of topological scalability and topological fault tolerance is defined. The dependence of the minimum of a topological fault tolerance by the girth of the system graph is shown. Model of parallel computings, and functions of the topological fault tolerance and scalability are adapted to the existence of unique nodes in information topology of the solved task. A method for configuring fault-tolerant subsystems for a deficient topological fault tolerance of a computing system is proposed, while providing the preassigned fault multiplicity for the solved task is achieved by duplicating subsystems which are configured for less, than the preassigned, fault multiplicity.

Keywords: scalable computing systems, their topological fault-tolerance.

*Статья представлена к публикации
членом редакционной коллегии Н.Н. Непейвода.*

*Поступила в редакцию 20.09.2016.
Опубликована 30.11.2017.*