

УДК 021.8 + 025.1
ББК 78.34

О ТОПОЛОГИЧЕСКОЙ МАСШТАБИРУЕМОСТИ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ¹

Мелентьев В. А.²

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

Большая часть опубликованных в научной литературе результатов, посвященных исследованию масштабирования параллельных задач и систем, локализованы под конкретные их реализации. Эксклюзивность, обусловленная различиями в классах решаемых задач, в техническом, технологическом и топологическом воплощении систем, во-первых, не позволяет напрямую использовать эти результаты в анализе и синтезе иных параллельных систем и задач, и во-вторых, не дает цельной картины взаимной обусловленности заданных параметров системы с востребованными показателями ее функционирования. На основе предложенной в настоящей работе модели предпринята попытка восполнить имеющийся в этом отношении пробел, и прежде всего, оценить влияние топологии на масштабируемость параллельных систем и решаемых на них задач.

Ключевые слова: топологии, сетевые технологии систем, масштабируемость параллельных вычислительных систем и задач.

1. Введение

Понятие масштабируемости может быть отнесено к любой системе как свойство, характеризующее зависимость критиче-

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

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

ски важных (существенных) показателей качества функционирования от числа составляющих ее конструктивно и/или функционально самостоятельных элементов. При этом потребности потенциальных пользователей системы определяют множество оцениваемых показателей, а составные элементы, способы их соединения и взаимодействия – вариативное множество «измерений», в которых следует оптимизировать эти показатели. Как правило, одним из основных показателей качества вычислительной системы (ВС) считается эффективность использования всей совокупности ее процессоров. Однако такая эффективность представляет интерес прежде всего для собственников системы. Соответствие системы потребностям конкретных пользователей оценивают эффективностью использования не всех, а только тех процессоров ВС, которые задействованы в образуемых для реализации соответствующих приложений подсистемах.

Достижение наивысших значений эффективности использования компонентов системы далеко не всегда является целью оптимизации ее архитектуры или прикладных алгоритмов. Например, общесистемным задачам непременно сопутствует требование оперативности их решения, а актуализация приложений является причиной создания специализированных систем; в коммерческих системах реального времени определяющим фактором также часто является оперативность пользовательских приложений, даже в ущерб эффективности использования ресурсов. Однако в этих случаях тоже можно говорить об эффективности, но уже не в части использования ресурсов, а в отношении повышения актуальности получаемых результатов.

Целью данной работы является формализация обусловленной топологией вычислительной системы зависимости ее потенциальных в отношении параллелизма возможностей при горизонтальном масштабировании¹. Естественно, что исследование такой топологической обусловленности предполагает абстрагирование от ограничений, связанных с присутствием в

¹ Под горизонтальным масштабированием ВС понимается наращивание ее возможностей увеличением числа вычислительных модулей.

параллельном алгоритме скалярных (нераспараллеливаемых) фрагментов, учитываемых классическим вариантом закона Амдала, названного в [5] «бичом параллельных вычислений». При исключении из рассмотрения фактора «скалярности» основным обстоятельством, определяющим реальную производительность вычислительной системы, становится ее зависимость от межпроцессных обменов, связанных как с топологией системной сети, так и с ее технологическими характеристиками – темпом выдачи сообщений, задержкой, пропускной способностью, и эта зависимость тем существеннее, чем больше размер системы [1].

В работе предложена модель параллельных вычислений, на базе которой проведено размежевание топологических и технологических факторов по их влиянию на реальную производительность системы и на предельный (при директивных критериях эффективности) порядок подсистем; даны формальные определения функций топологической масштабируемости параллельных задач и системы, основанные на использовании показателей плотности графа ВС с достижимостью, лимитированной используемой в системе сетевой технологией, параметрами задачи и заданными критериями эффективности их решения.

2. Описание модели масштабируемой ВС

Понятно, что оценка влияния сетевой топологии на потенциальную масштабируемость задач в масштабируемой ВС будет правомерной, если количественный критерий оценки такого влияния не меняет относительную упорядоченность сопоставляемых топологий при допускаемых предназначением ВС изменениях класса решаемых задач и обрабатываемых данных. В принципе, сетевая топология ВС может быть спроектирована под определенную задачу в смысле биективности ее информационно-логической структуры физической структуре системы – это позволяет добиться достаточно высоких показателей качества таких систем, но только при решении ими соответствующих специализации задач. Отметим также, что специализация систем и эффективность их использования в задачах, выходя-

щих за рамки специализации, антагонистичны по своей сути. Поэтому случай узкой специализации в данной работе мы не затрагиваем и далее будем вести речь о системах универсального назначения, пригодных для решения достаточно широкого класса задач и увеличивающих, таким образом, круг потенциальных потребителей и коммерческую эффективность их использования.

Известная в научной литературе концепция неограниченно-параллелизма, с вербальным описанием которой можно ознакомиться, например, в [3] и [4], вводит ряд допущений в свойства параллельной системы в целом, не разделяя при этом свойства распараллеливаемого приложения и свойства реализующей его вычислительной системы. Для выявления влияния на параллелизм общесистемных архитектурных особенностей ВС (в нашем случае – топологии) мы далее абстрагируемся от приложений, считая их неограниченно распараллеливаемыми при том, что все ограничения параллелизма обусловлены только архитектурой вычислительной системы.

Из обобщенного формального описания такой постановки выделим часть, отнесенную только к задаче. Здесь W и w – измеряемые временем объемы вычислений при решении произвольной задачи на одном и на p процессорах вычислительной системы. Соответствующие числу задействованных процессоров p объемы подлежащих обмену данных обозначим через Q и q , измеряя их при этом информационными единицами (байтами):

1. Задача допускает разбиение на произвольное число p информационно-связанных параллельных ветвей, $1 \leq p \leq \infty$. Информационный граф W_p 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 п.п. 1 и 3.

Понятно, что если бы в системе могла быть реализована полносвязная топология с реактивностью сетевой технологии¹, сравнимой с быстродействием процессора, то это соответствовало бы концепции неограниченного параллелизма системы в целом, и проблемы наращивания быстродействия ВС не возникало бы. Однако, во-первых, полная связность процессоров при неограниченном наращивании числа процессоров реализована быть не может как из экономических соображений, так и из-за отсутствия технической возможности неограниченного наращивания числа портов в процессорах. Во-вторых, из чисто физических ограничений (скорости света и несравнимости расстояний внутри процессора и между многими процессорами системы) даже предельное быстродействие NT не будет адекватным быстродействию процессора и его оперативной памяти, – к примеру, возможности высокоскоростной передачи данных по медным проводникам ограничены 40 Гбит/с, что соответствует единицам сантиметров на плате и примерно одному метру при передаче по кабелю [11]. Поэтому сдерживающим фактором линейного наращивания вычислительной мощности является наличие задержек в реализации межпроцессорных обменов. В связи с этим данный выше перечень свойств задачи с неограниченным параллелизмом дополним свойствами реализующей ее вычислительной системы, влияющими на величину таких задержек:

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 регулярной² и неполносвязной.

¹ Далее в тексте статьи словосочетание «сетевая технология» заменено аббревиатурой NT (*Network Technology*).

² Степени всех вершин графа ВС одинаковы.

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.

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

Перечисленные здесь свойства неограниченно распараллеливаемой задачи (п.п. 1-4) и свойства (п.п. 5-10) предназначенной для решения таких задач вычислительной системы в совокупности определяют используемую далее модель параллельных вычислений.

3. Технологический и топологический аспекты масштабируемости параллельных вычислений

Предпоследнее из свойств используемой нами модели параллельных вычислений является тривиальным в практике построения и использования коммуникационных сетей: как правило, время передачи сообщений между наиболее удаленными элементами оценивают диаметром соответствующих графов [10], характеризующим коммуникационную задержку в худшем случае [9]. Учитывая это, представим функцию $T_{ND}(p)$ сетевых задержек (ND – Network Delay) произведением определяемой топологией функции расстояния $L(p)$ между наиболее удаленными в графе ВС и задействованными в решении задачи процессорами на определяемую NT и зависящую¹ от $q = Q/p$ функцию задержки $t_{NT}(q)$: $T_{ND}(p) = L(p) \times t_{NT}(q)$. Отметим при этом, что неперенным свойством функции $t_{NT}(q)$ является ее обратная от p и прямая от Q зависимость: $p_1 < p_2 \Rightarrow t_{NT}(Q/p_1) > t_{NT}(Q/p_2)$ и $q_1 > q_2 \Rightarrow t_{NT}(q_1) > t_{NT}(q_2)$. Применение в ВС технологий с иной зависимостью функции $t_{NT}(q)$

¹ В соответствии с п. 3 используемой здесь концепции.

противоречило бы основной цели распараллеливания – достижению требуемой оперативности в реализации пользовательских приложений и требуемой их достоверности, связанной с повышением сложности алгоритмов и/или объемов обрабатываемых данных. Однако, к сожалению, увеличение числа процессоров p , приводящее к пропорциональному уменьшению удельного объема q обмениваемых данных, приводит к уменьшению элементарного (на единичном расстоянии) времени $t_{NT}(q)$ информационного взаимодействия физически смежных процессоров не в тех же, а в зависимых от применяемой в системе NT пропорциях, т.е. с некоторым коэффициентом непропорциональности k . К примеру, при коммутации сообщений функция элементарных задержек от числа процессоров p является гиперболической: $t_{NT}(q) = \alpha + q / \beta$, здесь α – латентность сети, β – ее пропускная способность [6], а $q = Q / p$ – объем обмениваемых данных между двумя информационно смежными процессорами. Из п. 3, в соответствии с которым $\lim q = 0$, и из определения латентности [12] как времени, затрачиваемого на передачу информационного сообщения с нулевым объемом данных, ясно, что ее значение является горизонтальной асимптотой в графике гиперболической функции $t_{NT}(p)$ при заданном объеме Q . Подобная этой зависимость $t_{NT}(p)$, хоть и в разной степени, но справедлива не только для рассмотренного здесь в качестве примера метода коммутации сообщений, но и для других используемых в современных сетях связи ВС сетевых технологий.

Рассмотрим общий случай масштабирования задачи, включающий в себя масштабирование числа p используемых ею процессоров с коэффициентом $k_p \geq 1$ и масштабирование этой задачи «по данным», увеличивающее общий объем обмениваемых данных Q в $m \geq 1$ раз. При этом элементарное время задержки

$$t_{NT} \left(\frac{m \cdot q}{k_p \cdot p} \right),$$

в соответствии с оговоренным выше свойством непропорционального его изменения и в зависимости от диапазона изменения обобщающего эти виды масштабирования коэффициента $k_{mp} = m / k_p$, определим неравенствами:

$$k_{mp} > 1 \Rightarrow t_{NT}(q) < t_{NT}(k_{mp} \cdot q) < k_{mp} \cdot t_{NT}(q),$$

$$(1) \quad k_{mp} = 1 \Rightarrow t_{NT}(q) = t_{NT}(k_{mp} \cdot q),$$

$$k_{mp} < 1 \Rightarrow k_{mp} \cdot t_{NT}(q) < t_{NT}(k_{mp} \cdot q) < t_{NT}(q).$$

Неравенства учитывают непропорциональность изменения времени задержки $t_{NT}(q)$, заданного общим объемом обмениваемых данных Q и числом задействованных в решении задачи процессоров p , от масштабирования как этого объема Q , так и числа процессоров p . Эта непропорциональность обусловлена сетевой технологией, например, для рассмотренной выше NT с коммутацией сообщений – обратной пропорциональностью от числа процессоров p элементарного объема q и присутствием во времени постоянной составляющей – латентности α . Введем в неравенства (1) коэффициент

$$k_{NT} = k_{mp} \cdot t_{NT}(q) / t_{NT}(k_{mp} \cdot q),$$

учитывающий присущую используемой в системе NT непропорциональность, назовем этот коэффициент технологическим и определим диапазоны его изменения при изменении k_{mp} :

$$k_{mp} > 1 \Rightarrow k_{NT} > 1,$$

$$(2) \quad k_{mp} = 1 \Rightarrow k_{NT} = 1,$$

$$k_{mp} < 1 \Rightarrow k_{NT} < 1.$$

Из данного выше формального определения технологического коэффициента k_{NT} видно, что его величина определена исходным (до масштабирования) элементарным объемом $q = Q / p$, и так как $k_{mp} > t_{NT}(q) / t_{NT}(k_{mp} \cdot q)$ – на эту непропорциональность мы обращали внимание выше, – то k_{NT} находится в прямой (но не в прямо пропорциональной) зависимости от масштабного коэффициента k_{mp} . К примеру, технологический коэффициент уже рассмотренной выше NT с коммутацией сообщений –

$$k_{NT} = \frac{k_{mp} \cdot t_{NT}(q)}{t_{NT}(k_{mp} \cdot q)} = \frac{k_{mp}(\alpha + q / \beta)}{\alpha + k_{mp} \cdot q / \beta} = \frac{k_{mp}(\alpha\beta + q)}{\alpha\beta + k_{mp} \cdot q}.$$

Однако далее мы абстрагируемся от исходных значений p и q и ограничимся использованием определенных выражением (2) качественных оценок k_{NT} и k_{mp} , тем не менее достаточных для

установления взаимозависимости топологической и технологической составляющей архитектуры ВС.

Учитывая вытекающее из п. 8 используемой нами модели неравенство $T_{ND}(p) > w$ и считая обмены максимально совмещенными с вычислениями, определим время фактических задержек разностью $T_{ND}(p) - w$. Модифицируя при этом известный закон Амдала, получим:

$$(3) S_p = \frac{W}{w + (L_S(p) \cdot t_{NT}(q) - w)} = \frac{W}{L_S(p) \cdot t_{NT}(q)}.$$

Тогда максимально допустимое при решении задачи с параметрами W и q расстояние $L_S(p) \geq 1$, обусловленное директивным (S_p , предписанным) при заданном числе процессоров p ускорением S_p , определится из

$$(4) L_S(p) = \frac{W}{S_p \cdot t_{NT}(q)}.$$

Отметим, что максимальное для системы с некоторой NT ускорение $S_p = W / t_{NT}(q)$ в решении задачи с объемом Q обмениваемых данных может быть получено при предельно допустимом расстоянии $L_S(p)$ между информационно смежными процессорами, равном единице. Такое могло бы быть реализовано только при использовании в ВС полносвязной топологии – в этом случае информационная топология задачи не имеет значения – или при использовании в ВС топологии, повторяющей информационную топологию задачи с точностью до изоморфизма. Из выражений (3) и (4), учитывая известную для используемой NT зависимость времени задержки t_{NT} от числа процессоров p , несложно определить и нижнее граничное значение этого числа $p_S(L)$ для заданного ускорения S_p при изменении L .

Если для пользователя ВС приоритетным является не ускорение S_p , а эффективность $E_p = S_p / p$ использования задействованных в решении задачи p процессоров, то

$$(5) E_p = \frac{W}{p \cdot L_E(p) \cdot t_{NT}(q)},$$

и предельное расстояние $L_E(p)$ для данной такой эффективности E_p составит

$$(6) L_E(p) = \frac{W}{p \cdot E_p \cdot t_{NT}(q)}.$$

Учитывая, что сохранение прежней эффективности $E_{p+\Delta p} = E_p$ с помощью уменьшения расстояния $L_E(p + \Delta p) < L_E(p)$, компенсирующего увеличение $p + \Delta p$, невозможно из-за относительной несоизмеримости изменения этих величин — $(p + \Delta p) / p > L_E(p) / L_E(p + \Delta p)$, становится понятным, что связанное с потребностью повышения оперативности в решении задач увеличение числа p параллельных ветвей неизбежно ведет к снижению эффективности E_p использования процессоров — это, как отмечено выше, обусловлено непропорциональным к увеличению p уменьшением $t_{NT}(Q/p)$: $\Delta p > 0 \Rightarrow \Rightarrow (p + \Delta p) \cdot t_{NT}(Q/(p + \Delta p)) > p \cdot t_{NT}(Q/p)$ с соответствующим изменением знаменателей в (5) и (6).

Факт снижения эффективности при масштабировании задач подтвержден многочисленными экспериментальными исследованиями (например, в [13]), теоретическое же его обоснование нетрудно увидеть из рассмотрения полносвязной топологии, где $\forall p > 1 L(p) = 1$. Таким образом, определенное из (5), (6) при требуемой эффективности E и при заданном предельном расстоянии L число ветвей $p_E(L)$ параллельной (W, Q) -задачи в системе, использующей сетевую технологию с характерной для нее функцией задержки $t_{NT}(Q/p)$, является верхним пределом распараллеливания.

Учитывая это и то, что нижний предел p_S распараллеливания обусловлен заданным ускорением S , делаем вывод о том, что распараллеливание задачи с требуемыми значениями S и E возможно только если при обусловленном сетевой технологией ВС расстоянии L между информационно смежными в ней процессорами выполняется неравенство $p_S(L) \leq p_E(L)$, а топологически обеспечиваемое при этом L число процессоров p находится между значениями p_S и p_E — $p_S(L) \leq p(L) \leq p_E(L)$.

В дополнение к вышесказанному еще раз обратим внимание на различную природу обусловленности расстояния L и числа процессоров p в (W, Q) -задаче: если предельное расстояние $L(p)$ при заданном p обусловлено функцией $t_{NT}(Q/p)$, т. е. используемой в ВС сетевой технологией, то предельное число $p(L)$ процессоров, которое может быть задействовано при допуске задачей расстояния L , обусловлено только топологически и

зависит от информационной топологии задачи и физической топологии системы.

Отмеченную здесь инвариантность обусловленных директивными значениями E_p и S_p предельных расстояний $L_E(p)$ и $L_S(p)$ к топологии не следует понимать как гарантированную произвольной топологией ВС обеспеченность этих расстояний при заданном p и при прочих равных условиях – равенстве объемов W вычислений и Q обмениваемых в задачах данных, равенстве критериев эффективности реализации этих задач и идентичности используемой в системах NT . Инвариантность к топологии расстояний $L_E(p)$ или $L_S(p)$ состоит в том, что, во-первых, они определены вне зависимости от топологии ВС, и, во-вторых, они определяют общее для любой топологии требование, заключающееся в следующем: успешное (в смысле обеспечения критериев эффективности, предписанных задаче с параметрами W и Q) распараллеливание на p процессоров любой задачи возможно лишь тогда, когда топология системы гарантирует хотя бы одно вложение (W, Q) -задачи, при котором расстояния между информационно смежными процессорами в соответствующем этой задаче подграфе ВС не превышают этих расстояний. Отметим здесь интуитивно понятную прямую зависимость успешности (в данном выше смысле) вложения от величины расстояния $L(p)$: большему допускаемому расстоянию соответствует и большая вероятность наличия в нем подграфов, допускающих удовлетворяющие заданным требованиям вложения. Это объяснимо хотя бы тем, что с увеличением предельно допускаемого расстояния возрастает число удовлетворяющих этому пределу вершин графа, естественно при этом, что кратность возрастания числа таких вершин должна быть выше кратности масштабирования k_p , но этот момент уже определяется топологиями задачи и системы.

Рассмотрим, каким образом технологическая составляющая ВС влияет на масштабирование задачи с коэффициентом $k_{mp} = m / k_p$, где коэффициент $m \geq 1$ соответствует увеличению объема Q в m раз, а $k_p \geq 1$ – увеличению числа задействованных в решении этой задачи процессоров в пределах n ($p \cdot k_p \leq n$), где n – общее число процессоров в системе (порядок графа ВС). Дополнительно к уже использованным ранее обозначениям

ускорения S_p , эффективности E_p и соответствующих этим критериям предельно допустимых при заданном числе процессоров p расстояний $L_E(p)$ и $L_S(p)$ введем обозначения ускорения S_{mp} , эффективности E_{mp} и соответствующих расстояний $L_{S_{mp}}(p)$ и $L_{E_{mp}}(p)$ при масштабированных с коэффициентом m данных и масштабированном с коэффициентом k_p исходном числе процессоров p . В соответствии с п. 2 изложенной в разделе 2 модели объем вычислений W при этом также увеличивается в m раз. Тогда из (6), оставив прежнюю (до масштабирования) эффективность $E_{mp} = E_p$ и используя введенный выше технологический коэффициент $k_{NT} = k_{mp} \cdot t_{NT}(q) / t_{NT}(k_{mp} \cdot q)$, получим

$$(7) \quad L_{E_{mp}}(p) = \frac{m \cdot W}{k_p \cdot p \cdot E_{mp} \cdot t_{NT}(k_{mp} \cdot q)} = \frac{W \cdot k_{NT}}{p \cdot E_p \cdot t_{NT}(q)} = k_{NT} \cdot L_E(p).$$

Если масштабирование числа задействованных в решении задачи процессоров превалирует над масштабированием данных и $k_p \geq m$, то в соответствии с (2) $k_{NT} < 1$, и из (7) – $L_{E_{mp}}(p) < L_E(p)$. Таким образом, сохранение эффективности при масштабировании задачи с кратностью увеличения объема данных m , меньшей кратности увеличения числа процессоров k_p , потребует обеспечения в k_{NT}^{-1} раз меньшего $L_E(p)$ предельно допустимого расстояния на в k_p раз большем множестве информационно смежных процессоров. Понятно, что явная противоречивость этого предъявляет более жесткие требования к топологии системы, каковой в условиях неограниченного (при $k_p \rightarrow \infty$) масштабирования может быть только полносвязная топология.

Если же кратность m масштабирования данных в задаче более или равна кратности масштабирования k_p числа задействованных при ее решении процессоров – $k_{mp} \geq 1$, то в соответствии с (2) $k_{NT} \geq 1$, при этом из (7) следует $L_{E_{mp}}(p) \geq L_E(p)$, и превышение $L_{E_{mp}}(p)$ над $L_E(p)$ будет тем большим, чем больше m превышает k_p . Как видим, здесь, в отличие от предыдущего случая, сохранению прежней эффективности при увеличении числа задействованных в решении задачи процессоров способствует возможность увеличения предельно допустимого расстояния между ними. Естественно, потенциал такого увеличения для единственной задачи определялся бы близостью (в лучшем случае – изоморфностью) топологии системы к топологии раз-

мещаемой в ней задачи. Для множества же решаемых в системе задач потребуется оптимизация топологии ВС под все задачи этого множества с учетом присущих им объемов вычислительных W и обменных Q операций, а также с учетом ожидаемых пределов масштабирования этих задач.

Из изложенного ясно, что сетевая технология играет существенную роль в организации эффективного функционирования ВС. Для каждой из решаемых в системе параллельных (W, Q)-задач, характеризующихся вычислительной сложностью W и общим объемом Q обменных взаимодействий, вне зависимости от их информационных топологий и от физической топологии ВС сетевая технология определяет потребную для достижения директивной эффективности решения задачи функциональную зависимость между числом задействованных процессоров p и предельно допустимым расстоянием $L(p)$ между ними. Отвечающая такой зависимости топология, позволяющая в рамках заданного набора из N решаемых в ВС задач с заданными рангами p_i ($i = 1, \dots, N$) их распараллеливания конфигурировать подсистемы с меньшими $L(p_i)$ расстояниями между информационно смежными процессорами, дает возможность снизить требования к технологической составляющей системы – для осознания этого факта достаточно рассмотреть предельный случай использования полносвязного графа ВС. Поэтому, учитывая, что отнесенное к i -й задаче ранга p_i расстояние $L(p_i)$ является предельно допускаемым в рассматриваемой параллельной системе, топологическую ее составляющую имеет смысл характеризовать адекватностью топологии ВС требованиям, предъявляемым одной или набором из N решаемых в ней задач в обеспечении соответствующих этим задачам предельных расстояний $L(p_i)$.

Таким образом, обусловленное директивной для отдельной задачи эффективностью предельное расстояние может быть обеспечено выбором соответствующей этой задаче топологии. Если же задач, определяющих предназначение системы, некоторое множество, то неизбежно встает вопрос о выборе топологии, оптимизированной под это множество. Однако даже при самой

лучшей оптимизации никакая, кроме полносвязной¹, топология при ее статической организации не способна обеспечить изоморфную вложимость всех топологически разнотипных² задач из набора, заданного ориентацией системы. Поэтому даже без учета предполагаемого в перспективе масштабирования задач необходимость организации опосредованных (транзитных) взаимодействий информационно смежных в задачах процессоров становится неизбежной, и расстояния между ними зависят от степени адекватности топологии задачи топологии системы. Масштабирование системы – увеличение общего числа процессоров n системы – приводит к такому же (в смысле увеличения расстояний) результату в связи со стремлением максимального их использования в решении той или иной задачи из набора (масштабирование задач). Кроме того, связанное с масштабированием задачи наращивание числа вершин ее информационного графа может сопровождаться ростом его степени и снижением вероятности изоморфного его вложения в граф ВС вплоть до нулевой при превышении степенью первого степени второго. Добавим к этому, что увеличение числа p задействованных в решении задачи процессоров неполиномиально увеличивают сложность и время реализации алгоритмов ее изоморфного размещения в системе, и при числе процессоров, большем десяти, это побуждает к замене точных методов приближенными, со свойственным им полярным отношением достоверности и актуальности.

Итак, потенциальные возможности масштабирования задачи, свободные от ограничений, вносимых используемыми в системе сетевой технологией и топологией, определяются общим числом n процессоров и возможностью наращивания их числа. Существенная, в сравнении с вычислительными операциями, продолжительность информационных взаимодействий ограничивает этот потенциал: число p задействованных в решении задачи процессоров не может быть вне диапазона значений,

¹ В полносвязной ВС расстояние между любыми двумя вершинами не зависит от числа процессоров и является единичным.

² Здесь имеется в виду, что, информационно-логическая структура таких задач представлена существенно отличающимися графами.

определенных из (4), (6) и соответствующих директивным значениям ускорения и эффективности. Для узкоспециализированных ВС проблема адекватного вложения единственной задачи может быть решена выбором соответствующей этой задаче топологии. Расширение набора топологически разнотипных задач, определяющих проблемную ориентацию ВС, и потребности обеспечения возможностей эффективного их масштабирования в наращиваемой системе не укладываются в рамки какой-либо одной, кроме полносвязной, топологии. Экономическая нецелесообразность и техническая неосуществимость реализации такой топологии ограничивают возможности эффективного масштабирования как отдельных решаемых в системе задач, так и системы в целом.

В связи с теоретической и физической невозможностью реализации в ВС полносвязной топологии, в работе [7] впервые представлено решение проблемы вложения задач в постановке, основанной на замене отношений смежности вершин графа ВС отношениями лимитированной их достижимости – при этом расстояния между вершинами графа подсистем, образуемых для решения распараллеленных на p ветвей задач, ограничены определяемыми из (4), (6) предельно допускаемыми расстояниями.

4. Вложение параллельных задач с учетом лимитированной достижимости

Как показано в предшествующем разделе, соотношение между объемами вычислений W и информационных взаимодействий Q в параллельной задаче, заданные критерии эффективности ее решения и быстродействие используемой в системе NT определяют взаимозависимость числа p задействованных процессоров и предельно допускаемого при этом p расстояния $L(p)$ между информационно смежными вершинами соответствующего задаче подграфа в графе ВС. Учитывая, что расстояния между вершинами невзвешенного графа $G(V, E)$ определяются числом транзитных участков и могут быть выражены только целыми числами, а также тем, что расстояние между информационно смежными вершинами подграфа не должно превышать

определяемого из (4) и/или (6) предельного при предписанных критериях S_p или E_p и при заданном p значения $L(p)$, предельное расстояние между информационно смежными в задаче процессорами определим целой частью $L(p)^1$ и назовем предельной достижимостью $\hat{\partial}(p)$:

$$(8) \quad \hat{\partial}(p) = \lfloor L(p) \rfloor, \quad 1 \leq \hat{\partial}(p).$$

В соответствии с (4) директивное ускорение S для уменьшенного в сравнении с $L_S(p)$ значения предельной достижимости $\hat{\partial}(p) \leq L(p)$ может быть достигнуто при большей задержке t_{NT} и, соответственно, при уменьшенном в сравнении с p числе процессоров $p_{\hat{\partial}}$. Покажем это на уже использованном примере NT с коммутацией сообщений, обозначив допускаемые при равенстве ускорений $S_p = S_{p_{\hat{\partial}}}$ задержки на соответствующих $L_S(p)$ и $\hat{\partial}(p)$ расстояниях, через $t_{NT}(q)$ и $t_{NT}(q')$. Из

$$t_{NT}(q) = \alpha + \frac{Q}{p \cdot \beta} \text{ и}$$

$$t_{NT}(q') = \alpha + \frac{Q}{p_{\hat{\partial}} \cdot \beta}$$

и из (4) получим минимальные при этих задержках значения числа процессоров:

$$p = \frac{Q \cdot L_S(p) \cdot S_p}{\beta \cdot (W - \alpha \cdot L_S(p) \cdot S_p)} \text{ и}$$

$$p_{\hat{\partial}} = \frac{Q \cdot \hat{\partial}(p) \cdot S_p}{\beta \cdot (W - \alpha \cdot \hat{\partial}(p) \cdot S_p)}.$$

Отсюда нетрудно убедиться в том, что $p_{\hat{\partial}} < p$.

Вершины u и v графа вычислительной системы $G(V, E)$ считаем связанными отношением $\hat{\partial}$ -достижимости² $R_{\hat{\partial}}$ ($\hat{\partial}$ -достижимыми), если расстояние $d(u - v)$ между ними не превышает $\hat{\partial}$: $(u - v) \in R_{\hat{\partial}} \Leftrightarrow d(u - v) \leq \hat{\partial}$. Порожденный множеством вершин $V_{\hat{\partial}} = V$ и матрицей $\hat{\partial}$ -достижимости $R_{\hat{\partial}}$ граф $G_{\hat{\partial}}(V_{\hat{\partial}}, E_{\hat{\partial}})$, $E_{\hat{\partial}} \in R_{\hat{\partial}}$ назван графом $\hat{\partial}$ -достижимости. Очевидно, что превышение степени этого графа над степенью графа $G(V, E)$ тем значительнее, чем больше полученное из (4) или из (6) и из $\hat{\partial}(p) = \lfloor L(p) \rfloor$ значение $\hat{\partial}(p)$. Максимальное по включению подмножество $V_{\hat{\partial}}^* \subseteq V_{\hat{\partial}}$ связанных отношением

¹ Здесь индекс критерия S_p или E_p опущен.

² Здесь и далее зависимость значения достижимости $\hat{\partial}$ от числа процессоров p не указана и предполагается по умолчанию $\hat{\partial} \equiv \hat{\partial}(p)$.

∂ -достижимости R_∂ вершин порождает подграф $G_\partial(V_\partial^*, R_\partial)$, являющийся кликой графа $G_\partial(V_\partial, E_\partial)$, иначе – ∂ -кликкой $K_\partial(G)$ графа $G(V, E)$. Подобно плотности $\varphi(G)$ графа G , определяемой порядком его наибольшей клики [2], порядок наибольшей из максимальных ∂ -клик графа $G(V, E)$ назовем плотностью $\varphi(G_\partial)$ графа ∂ -достижимости, или ∂ -плотностью $\varphi_\partial(G)$ графа $G(V, E)$ – $\varphi(G_\partial) \equiv \varphi_\partial(G)$.

В работе [7] введено понятие ∂ -ограниченной компоненты достижимости (∂ -компоненты) графа. Говоря далее о ∂ -компоненте графа $G(V, E)$, будем иметь в виду данное там определение глобальной компоненты ∂ -достижимости, как подграфа $G^*_{\partial+}(V^*_{\partial+}, E^*_{\partial+})$, порожденного максимальным по включению множеством $V^*_{\partial+} = V^*_{\partial} \cup V^*_+$ вершин, объединяющим подмножество $V^*_{\partial} = \{u \mid \forall v \in V_\partial(u - v) \in R_\partial\}$ попарно ∂ -достижимых вершин с подмножеством $V^*_+ = \{w \notin V^*_{\partial} \mid \exists u, v \in V^*_{\partial} : w \in V_\partial(u - v)\}$ вершин, не являющихся членами подмножества V^*_{∂} , но входящих в состав ∂ -ограниченных путей между его вершинами; здесь $V_\partial(u - v)$ – подмножество вершин, составляющих путь из u в v . Вершины из V^*_{∂} в плане их попарной ∂ -достижимости называем основными; они составляют множество вершин ∂ -клики $K_\partial(G)$ графа $G(V, E)$, и их число (порядок ∂ -клики, порядок глобальной ∂ -компоненты) есть ∂ -плотность $\varphi_\partial(G)$ графа $G(V, E)$ – $\varphi_\partial(G) = |V^*_{\partial}|$. Вершины из V^*_+ – вспомогательные; они обеспечивают ∂ -достижимость основных вершин из V^*_{∂} и, не входя в V^*_{∂} , входят в состав ∂ -путей между ними.

Проблема выявления в графе ВС $G(V, E)$ компоненты ∂ -достижимости (∂ -клики графа G) решена в [7] с помощью проективного описания графа [8]. Суть соответствующего алгоритма состоит в следующем:

1. Строим систему $n = |V|$ ∂ -уровневых проекций $P_\partial(v_i)$ графа $G(V, E)$, определяющих граф ∂ -достижимости $G_\partial(V_\partial, R_\partial)$.

2. В проекциях $P_\partial(v_i)$ графа $G_\partial(V_\partial, R_\partial)$ выявляем подмножества $\bar{V}_\partial(v_i)$ вершин, ∂ -недостижимых из соответствующих ракурсных вершин v_i : $\bar{V}_\partial(v_i) = V \setminus V_\partial(v_i)$.

3. Используя полученные в п. 2 подмножества $\bar{V}_\partial(v_i)$, строим матрицу \bar{R}_∂ смежности графа ∂ -недостижимости $\bar{G}_\partial(V, \bar{R}_\partial)$ и

выявляем в нем наибольшее множество независимых вершин (МНВ). Входящие состав полученного МНВ вершины образуют искомое подмножество V_{∂}^* вершин, порождающих клику $K_{\partial}(G)$ графа ∂ -достижимости $G_{\partial}(V_{\partial}, R_{\partial})$.

4. Проекции $P_{\partial+}^*$ искомой компоненты ∂ -достижимости $G_{\partial+}^*(V_{\partial+}^*, E_{\partial+}^*)$ получаем последовательным (от $j = \partial$ до $j = 1$) исключением из каждого j -го уровня проекций $\{P_{\partial}(v_i) \mid v_i \in V_{\partial}^*\}$ вершин, не принадлежащих V_{∂}^* и не имеющих продолжения на $(j + 1)$ -м уровне проекции, полученной на предыдущем шаге. Оставшиеся при этом в результирующих проекциях $P_{\partial+}^*$ вершины не принадлежат V_{∂}^* и являются дополнительными.

5. О топологической масштабируемости параллельных задач и вычислительных систем

Уточним используемые в этом разделе обозначения:

p – ранг рассматриваемой задачи (число параллельных ветвей задачи, число задействованных в решении задачи процессоров);

W_p – информационно-логический граф задачи, соответствующий ее распараллеливанию на p параллельных ветвей;

$\partial \equiv \partial(p) = \lfloor L(p) \rfloor$ – обусловленное числом процессоров p и требуемым критерием эффективности S_p или E_p предельно допустимое в графе ВС значение достижимости;

$\varphi(W_p)$ – плотность распараллеленного на p ветвей информационного графа задачи, порядок наибольшей клики в графе W_p ;

$\varphi_{\partial}(G)$ – ∂ -плотность графа $G(V, E)$, она же – порядок ∂ -клики $K_{\partial}(G)$ графа G , или порядок клики $K(G_{\partial})$ графа ∂ -достижимости $G_{\partial}(V_{\partial}, R_{\partial})$.

5.1. ТОПОЛОГИЧЕСКАЯ МАСШТАБИРУЕМОСТЬ ЗАДАЧИ

Понятно, что, если числу p ветвей (W, Q) -задачи, решаемой в рассматриваемой ВС¹, соответствует предельная достижи-

¹ Здесь и далее, говоря о ВС, по умолчанию предполагаем использование ею сетевой технологии, которая для (W, Q) -задачи ранга p с

мость $\delta(p)$, а порядок p информационного графа W_p вкладываемой подсистемы не выше δ -плотности $\varphi_\delta(G)$ графа G – $\varphi_\delta(G) \geq p$, то вложение графа W_p в эту δ -клику может быть произвольным, даже если граф W_p является полным (при этом $\varphi(W_p) = p$). Попытка наращивания при этом числа процессоров сверх определенного для графа G значения δ -плотности ($p + \Delta p > \varphi_\delta(G)$) неизбежно потребует размещения Δp вершин вне клики $K(G_\delta)$, что увеличит расстояния между информационно смежными в задаче вершинами как минимум на единицу: $\delta(p + \Delta p) - \delta(p) \geq 1$. Так как по определению $\delta(p) = \lfloor L(p) \rfloor$ и $\delta(p + \Delta p) = \lfloor L(p + \Delta p) \rfloor$, то при таком увеличении p предельно-допустимое расстояние $L(p)$ будет превышено – $L(p + \Delta p) - L(p) \geq 1$, что воспрепятствует сохранению директивного критерия эффективности.

В отличие от рассмотренного выше случая информационно-полносвязных задач, вложение в граф δ -достижимости G_δ графа G задач с $\varphi(W_p) < p$ (неполносвязных задач) допускает увеличение числа p процессоров сверх δ -плотности $\varphi_\delta(G) - \varphi(W_p) \leq \varphi_\delta(G) < p$, если в графе G_δ найдется подграф, изоморфный графу W_p задачи. К примеру, если исходная ВС является узкоспециализированной и ориентирована на решение некоторой задачи (или нескольких топологически однотипных задач) с неполносвязным ($p > \varphi(W_p)$) информационным графом W_p , то, в принципе, для этой ВС может быть выбрана топология, повторяющая информационную топологию задачи с точностью до изоморфизма. Если при этом допустимы опосредованные (с $\delta(p) > 1$) взаимодействия информационно смежных процессоров, то топология ВС может быть сведена к одной из известных регулярных топологий, граф δ -достижимости G_δ которой изоморфен графу W_p . Этого можно достичь, например, определив граф G системы суграфом¹ информационного графа W_p задачи, получаемым путем последовательного удаления из W_p ребер до сохраняющего изоморфизм $G_\delta \simeq W_p$ их минимума. Такой подход позволит минимизировать топологию проектируемой ВС, и

требуемыми значениями показателей эффективности ее решения обуславливает предельно допустимую достижимость δ .

¹ Суграф – часть графа, имеющая то же множество вершин, что и сам граф.

он может быть применен и для некоторого набора из N решаемых в системе задач. Соответствующую постановку, состоящую в объединении графов W_{p_i} задач из этого набора ($i = 1, \dots, N$), несложно представить, поэтому останавливаться на ней не имеет смысла.

Итак, связанные с масштабированием параллельной задачи изменения информационной ее топологии существенно влияют на возможности вложения в ВС этой задачи, релевантного требуемым ускорению и эффективности. Понятно, что возможности вложения информационного графа W_p в граф δ -достижимости G_δ при превышении рангом p задачи δ -плотности $\varphi_\delta(G)$ графа ВС коррелированы не только информационной топологией задачи, но и физической топологией системы, и шансы изоморфного вложения задачи ранга p в граф δ -достижимости G_δ возрастают как с уменьшением плотности $\varphi(W_p)$ вкладываемых задач, так и с увеличением δ -плотности $\varphi_\delta(G)$ графа ВС (естественно при этом, что порядок p информационного графа W_p задачи не должен превышать порядка n графа G системы: $p \leq n(G)$). Очевидно, что тезис об увеличении шансов успешного вложения для задач с меньшей плотностью информационного графа актуален для любой системы, топологически адекватной заданному неравенству $\varphi(W_p) \leq \varphi_\delta(G)$ условию. Это позволяет абстрагироваться от используемой в ВС топологии (разумеется, в рамках ее адекватности этому условию) и, таким образом, сосредоточиться только на топологическом аспекте масштабируемости $\mu_\varphi(W_p)$ параллельной задачи, оценивая связанное с наращиванием числа параллельных ветвей относительное изменение возможностей ее вложения лишь качественно:

$$(9) \quad \mu_\varphi(W_p) = 1 - \frac{\varphi(W_p)}{p} = \frac{p - \varphi(W_p)}{p}.$$

Из (9) ясно, что в случаях, например, с неизменной при масштабировании задачи плотностью $\varphi(W_p) = 2$ и с плотностью $\varphi(W_p) = p$ (эти случаи характерны для конвейерных ($W_p \equiv C_p$) и полносвязных ($W_p \equiv K_p$) вычислений) для предельных значений $p = 2$ и $p = \infty$ получим, соответственно, $\mu_\varphi(C_2) = 1/2$, $\mu_\varphi(K_2) = 0$ и $\mu_\varphi(C_\infty) = 1$ и $\mu_\varphi(K_\infty) = 0$. Физический смысл введенной здесь функции $\mu_\varphi(W_p)$ состоит в том, что в первом случае наращивание параллелизма (числа p ее параллельных ветвей) уменьшает

относительную (в отношении к p) плотность информационного графа задачи, соответственно, его вложение в граф δ -достижимости G_δ , $1 \leq \delta \leq d(G)$, облегчается, и масштабируемость $\mu_\varphi(C_p)$ стремится к равному единице максимальному значению. Во втором же случае $\varphi(K_p) = p$, т.е. вне зависимости от ранга p относительная плотность параллельной задачи¹ максимальна и равна единице. Это, как и нулевое значение $\mu_\varphi(K_p)$, говорит о максимальной сложности масштабирования информационно полносвязной параллельной задачи и о том, что вложение в граф δ -достижимости² G_δ задачи K_p возможно, только если плотность $\varphi(G_\delta)$ этого графа не меньше числа ветвей p .

Однако относительная плотность – не единственное, что определяет сложность вложения масштабируемой задачи. К примеру, задачи одного ранга с кольцевой и звездной информационными топологиями обладают одинаковой плотностью, но степень s «кольца» C_p при этом независимо от ранга p задачи постоянна и $s(C_p) = 2$, тогда как наращивание числа p ветвей в задаче со «звездной» ($W_p \equiv Z_p$) топологией увеличивает степень ее информационного графа: $s(Z_p) = p - 1$, и при $p > s(G_\delta) + 1$ одно из необходимых условий наличия изоморфных вложений $s(Z_p) \leq s(G_\delta)$ будет нарушено. Коррелированную степень информационного графа масштабируемость задачи с аналогичным вышеописанному физическим смыслом можно описать тем же выражением, что и в (9), заменив плотность $\varphi(W_p)$ степенью $s(W_p)$:

$$(10) \mu_s(W_p) = 1 - \frac{s(W_p)}{p} = \frac{p - s(W_p)}{p}.$$

Тогда для задач с «кольцевой» и «звездной» топологиями при $p > 2$ получим: $\mu_s(C_p) = 1 - 2p^{-1}$ и $\mu_s(Z_p) = p^{-1}$, что соответствует увеличению $\mu_s(C_p)$ (улучшению возможностей вложения при наращивании «кольца») и уменьшению $\mu_s(Z_p)$ (ухудшению таких возможностей для «звезды»).

Совместное использование введенных выше частных функций масштабируемости параллельной задачи, связанных с изме-

¹ Словосочетание «информационного графа» для краткости опущено.

² Предельная достижимость δ соответствует заданному числу ветвей p .

нением плотности и степени ее информационного графа, позволяет качественно (больше/меньше – лучше/хуже) сопоставлять возможности вложения задач в регулярные топологии вычислительных систем, предельные значения достижимости δ в которых адекватны изменениям p , а плотность $\varphi(G_\delta)$ и степень $s(G_\delta)$ соответствующих графов δ -достижимости удовлетворяют условиям: $\varphi(G_\delta) \geq \varphi(W_p)$ и $s(G_\delta) \geq s(W_p)$. Функцию $\mu(W_p)$ зависимости топологической масштабируемости задачи от числа p параллельных ее ветвей определим произведением функций (9) и (10): (11) $\mu(W_p) = \mu_\varphi(W_p) \cdot \mu_s(W_p)$.

Из введенных выше частных функций (9), (10) и из обобщающей их функции (11) видим, что топологическая масштабируемость задачи ухудшается с приближением к единице относительных значений плотности или степени ее информационного графа, т.е. с приближением ее информационной топологии к полностью связной. Понятно также, что при сопоставлении алгоритмов распараллеливания задачи для обеспечения хорошей ее масштабируемости следует отдать предпочтение тем из них, в которых увеличение параллелизма приводит по крайней мере к меньшему росту относительных значений плотности и степени информационного графа.

5.2. ТОПОЛОГИЧЕСКАЯ МАСШТАБИРУЕМОСТЬ СИСТЕМЫ

Понятно, что степень и плотность графа δ -достижимости $G_\delta(V, E_\delta)$, $\delta > 1$, положительно коррелированы степени исходного регулярного¹ графа $G(V, E)$ и величиной полученного из (8) значения достижимости $\delta(p) = \lfloor L(p) \rfloor$: для графов G и H одного порядка ($n(G) = n(H)$)

$$s(G) < s(H) \Rightarrow s(G_\delta) < s(H_\delta), \quad \varphi_\delta(G) < \varphi_\delta(H),$$

$$\delta_1 < \delta_2 \Rightarrow s(G_{\delta_1}) < s(G_{\delta_2}), \quad \varphi(G_{\delta_1}) < \varphi(G_{\delta_2}).$$

Понятно также, что δ -плотность $\varphi_\delta(G)$ графа ВС как порядок $n(K_\delta)$ его максимальной δ -клик, все вершины которой по определению взаимно δ -достижимы, ограничивает сверху параллелизм (число ветвей) только для информационно полностью связных

¹ Напомним, что в соответствии с п. 6 рассматриваемой здесь модели топология ВС представлена регулярным неполностью связным графом.

задач. Для задач же с плотностью информационного графа $\varphi(W_p)$, меньшей числа ветвей p в них, верхние границы могут быть большими δ -плотности графа G ВС, если в графе δ -достижимости G_δ найдется подграф порядка $p > \varphi_\delta(G)$, изоморфный графу W_p . К примеру, число задействованных в гиперкубической ВС процессоров для задачи с кольцевой топологией ограничено сверху порядком гиперкуба вне зависимости от допускаемой при ее решении достижимости δ .

Итак, пределы распараллеливания решаемых в системе задач в зависимости от допускаемой для каждой из них достижимости δ находятся в диапазонах от $\varphi_\delta(G)$ до $n(G)$, поэтому задачи можно классифицировать соответственно принадлежности этим диапазонам, идентифицируемым значениями $\delta \geq 1$. Понятно при этом, что при использовании в системе достаточно быстродействующей NT допускаемая некоторыми задачами предельная достижимость δ не ограничивается диаметром $d(G)$ графа ВС, и $\delta \geq d(G) \Rightarrow \varphi_\delta(G) = n(G_\delta)$.

Рассмотрим, к примеру, решение двух из множества принадлежащих δ -классу задач на системах, топологии которых не привязаны к информационным топологиям этих задач. Пусть топологии систем заданы графами G_1 и G_2 одного порядка $n(G_1) = n(G_2)$, значения их δ -плотности соответствуют оговариваемому δ -классу, но различны: $\varphi_\delta(G_2) > \varphi_\delta(G_1)$. Наделим рассматриваемые здесь задачи W_1 и W_2 значениями топологической масштабируемости μ_1 и μ_2 , близкими к граничным и равными, соответственно, единичному и нулевому значениям. Следует ожидать при этом, что предел распараллеливания первой, наименее топологически сложной для вложения задачи для обеих систем будет близок к одному и тому же значению – порядку n их графов. Для второй задачи, как наиболее топологически сложной для вложения, можно ожидать, что предельное число задействованных в каждой системе процессоров будет близким к порядку наибольшей δ -клики графа – значениям δ -плотности $\varphi_\delta(G_1)$ и $\varphi_\delta(G_2)$. Из этого ясно, что системы с меньшей δ -плотностью графов обладают меньшими возможностями распараллеливания прежде всего топологически более сложных

задач, обладающих меньшей топологической масштабируемостью $\mu(W_p)$.

Понятно тогда, что при решении набора задач усредненное изменение пределов их распараллеливания будет более выигрышным в системах с большими значениями δ -плотности их графов. Поэтому вполне оправданным будет использование δ -плотности в качественной (больше/меньше – лучше/хуже) характеристике графов ВС и в сопоставительном их анализе не только для информационно полносвязных задач, но и для задач, не требующих информационной полносвязности. Естественным при этом будет нормирование значений δ -плотностей сопоставляемых графов ВС порядком этих графов — отношением $\varphi_\delta(G)/n$. Нормированное таким образом значение не только характеризует в абсолютном выражении потенциал системы в распараллеливании информационно полносвязных задач с $\mu(W_p) = 0$, но и позволяет осуществлять качественную оценку потенциала для классифицированных по достижимости $\delta > 1$ задач с $\mu(W_p) > 0$. Эту функцию изменения δ -плотности графа ВС от его порядка n $\mu_\delta(G_n) = \varphi_\delta(G_n)/n$ назовем функцией топологической масштабируемости системы.

При этом, как мы уже говорили выше, если рассматриваемая система допускает для некоторых из решаемых на ней задач $\delta \geq d(G)$, то предел распараллеливания этих задач определяется числом процессоров в ней — $\varphi_\delta(G) = n(G_\delta)$, понятно, что масштабируемость таких задач в этой системе максимальна и равна единице — $\mu_\delta(G_n) = 1$. Отсюда, кстати, вытекает желательность выбора сетевой технологии масштабируемой системы с учетом того, чтобы быстродействие NT было достаточным для эффективного использования увеличенного в результате предполагаемого в процессе эксплуатации масштабирования системы числа ее процессоров: например, для ключевых задач, определяющих функциональное назначение системы, обусловленная сетевой технологией достижимость δ должна быть близкой к текущему диаметру $d(G)$ графа ВС или превышать его.

Итак, учитывая, что предельные значения достижимостей δ в системе обусловлены предписанными ей задачами и используемой сетевой технологией, топологическую масштабируемость

системы, как характеристику изменения потенциала в распараллеливании предписанных ей задач при изменении порядка n графа G_n системы можно характеризовать семейством функций $\mu_{\hat{c}}(G_n)$ относительно к масштабу системы изменения \hat{c} -плотности графа G_n , описывающего топологию исследуемой ВС, при заданных значениях достижимости $\hat{c} > 1$.

6. Заключение

Все параллельные системы по мере их появления и наращивания подвергаются исследованиям их быстродействия и эффективности на различных классах и наборах задач и данных. Понятно, что различия в технической, технологической, топологической, прикладной и прочих архитектурных составляющих системы придают полученным при этом результатам некую эксклюзивность и могут быть распространены на другие системы и задачи лишь с некоторыми ограничениями. Обобщение полученного таким образом и опубликованного к настоящему времени огромного фактологического материала убеждает в качественной идентичности локального влияния этих компонент на показатели масштабируемых систем, но не дает цельной картины взаимной их обусловленности, достаточной для получения требуемых показателей вновь проектируемых и уже эксплуатируемых систем. В настоящей работе предпринята попытка восполнить имеющийся в этом отношении пробел, прежде всего, в оценке влияния топологии на масштабируемость параллельных систем и решаемых на них задач.

С этой целью во втором разделе статьи предложена разделенная на две составляющие модель параллельных вычислений: первая отнесена к параллельным приложениям и приписывает им свойства неограниченной распараллеливаемости, вторая отнесена к вычислительной системе, ограничения параллелизма которой обусловлены недостаточным быстродействием ее коммуникационной среды. В связи с тем, что задержки информационных взаимодействий ветвей зависят от присущих каждой задаче объемов вычислительных и обменных операций, естественной является характеристика задач соответствующими этим объемам параметрами.

Использование предложенной модели позволило сконцентрировать внимание на технологической и топологической составляющих коммуникационной среды и исследовать формальную их взаимозависимость в совокупном влиянии на пределы распараллеливания в заданной ВС задач, соответствующие предписанным критериям эффективности их решения. В третьем разделе статьи получены выражения, связывающие с используемой в системе сетевой технологией предельно допускаемые между информационно смежными процессорами расстояния $L(p)$, минимально необходимое и максимально допускаемое число параллельных ветвей, релевантные заданному ускорению и эффективности. Это позволяет оценить любой из перечисленных выше параметров при заданных остальных.

В связи с тем, что расстояния между вершинами невзвешенного графа определяются числом транзитных участков и могут быть выражены только целыми числами и не должны превышать при этом упомянутых выше предельно допускаемых расстояний между информационно смежными в задаче вершинами графа ВС, введено понятие предельной достижимости $\delta(p)$, определенной целой частью $L(p)$. В четвертом разделе рассматривается проблема вложения в граф ВС параллельных задач с учетом лимитированной достижимости. Даны понятия графа δ -достижимости, δ -компоненты графа, представляющей собой клику графа δ -достижимости и соответствующей порядку последней, δ -плотности графа ВС, приведен алгоритм выявления в графе ВС такой клики.

В пятом разделе введены взаимно абстрагированные формализованные показатели и функции топологической масштабируемости параллельных задач и топологической масштабируемости систем. Свойство взаимной абстрагированности, с одной стороны, позволяет, не привязываясь к конкретной среде реализации задачи, выбрать наименее топологически сложный для ее вложения алгоритм, а с другой – выбирать для унифицированных систем топологии, при прочих равных условиях обладающие наибольшими возможностями успешного вложения произвольных задач. Это не исключает, однако, использование данных функций для специализированных систем.

Введенные показатели топологической масштабируемости задач и систем основаны на использовании плотности информационного графа задачи и δ -плотности графа ВС; допускаемая предложенной в работе моделью нерегулярность топологии задачи учтена дополнением соответствующего показателя степени изменяющегося при ее масштабировании информационного графа. Отметим, что использование в составе показателей масштабируемости задач и вычислительных систем инвариантов соответствующих графов (степени, плотности и δ -плотности) определяет инвариантность этих показателей к способам вложения задач в системы, характеризуя, таким образом, потенциал масштабируемости, который в той или иной степени достижим при использовании этих способов.

Результаты работы будут полезными не только при анализе уже действующих систем и реализуемых на них параллельных алгоритмов, но и при создании новых систем или алгоритмов с учетом предполагаемого масштабирования тех и других.

Литература

1. АБРАМОВ С.М., ЛИЛИТКО Е.П. *Состояние и перспективы развития вычислительных систем сверхвысокой производительности // Информационные технологии и вычислительные системы.* – 2013. – №2. – С. 6–22.
2. БАДЕХА И.А., РОЛДУГИН П.В. *О плотности графов, в которых каждое ребро входит хотя бы в две максимальные клики // Дискрет. матем.* – 2013. – Т. 25, №3. – С. 7–21.
3. БУРОВА И.Г., ДЕМЬЯНОВИЧ Ю.К. *Алгоритмы параллельных вычислений и программирование: Курс лекций.* – СПб: Изд-во С.-Пб. ун-та, 2007. – 206 с.
4. ВОЕВОДИН В.В., ВОЕВОДИН В.В. *Параллельные вычисления.* – Санкт-Петербург. 2002. – 599 с.
5. ВОЕВОДИН В. *Суперкомпьютеры и парадоксы неэффективности // Открытые системы. СУБД.* – 2009. – №10. – С. 17–20.
6. ГЕРГЕЛЬ В.П. *Высокопроизводительные вычисления для многопроцессорных многоядерных систем: [учебник для ву-*

- зов]. – Нижний Новгород: Издательство Нижегородского госуниверситета, 2010. – 539 с.
7. МЕЛЕНТЬЕВ В.А. *Вложение подсистем, лимитирующих длину и число путей между вершинами графа вычислительной системы* // Управление большими системами. – 2014. – №47. – С. 212–246.
 8. МЕЛЕНТЬЕВ В.А. *Формальные основы скобочных образов в теории графов* // Труды II Международной конференции «Параллельные вычисления и задачи управления». РАСО'2004 Москва. 2004. – С. 694–706.
 9. ПОЖИЛОВ И.А. *Оценка задержки и пропускной способности сети с топологией «многомерный тор» при наличии отказавших каналов связи* // Научный сервис в сети интернет: многообразие суперкомпьютерных миров. Труды Международной суперкомпьютерной конференции. – С. 211–217.
 10. РАППОПОРТ А.М. *Метрические характеристики графов сетей коммуникаций* // Труды ИСА РАН. – 2005. – Т. 14. – С. 141–147.
 11. СЛЕПУХИН А.Ф. *Перспективы развития аппаратных технологий и их применение в суперкомпьютерах экзафлопного уровня* // Тезисы докладов Четвертого Московского суперкомпьютерного форума (Москва, 23 октября 2013 г.) / [Под. ред. Волкова Д.В. Москва. Россия. Октябрь 2013. – С. 7.
 12. ШПАКОВСКИЙ Г.И., СТЕЦЮРЕНКО В.И., ВЕРХОТУРОВ А.Е. И ДР. *Применение технологии MPI в Грид.* – Минск: Белорусский государственный университет, 2008. – 137 с.
 13. GRAMA A.Y., GUPTA A., KUMAR V. *Isoefficiency: Measuring the Scalability of Parallel Algorithms and Architectures* // Parallel & Distributed Technology: Systems & Applications, IEEE. – August 1993. – Vol. 1. No.3. – P. 12–21.

ON TOPOLOGICAL SCALABILITY OF COMPUTING SYSTEMS

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

Abstract: Most results on parallel task and system scaling met in the literature are guided by the specific technological and topological implementation. Unique properties of a problem originated from the class of the task solved, from technical, technological and topological implementation of the system, firstly, prevent using these results directly in the analysis and synthesis of the other parallel systems and tasks, and, secondly, hide the general pattern of mutual conditionality of the given parameters of system with demanded performance metrics. On the basis of the model suggested in this paper we try to fill the existing gap with the main goal to assess the impact of the topology of a parallel system on scalability of a system or a task.

Keywords: topology, system network technology, scalability of parallel computing systems and tasks.

Статья представлена к публикации членом редакционной коллегии М.В. Губко

*Поступила в редакцию 06.09.2015.
Опубликована 30.11.2015.*