

Теоретико-информационный подход к оценке производительности суперкомпьютеров

А. А. РАКИТСКИЙ^{1,2,*}, Б. Я. РЯБКО^{2,3}

¹Сибирский государственный университет телекоммуникаций и информатики, Новосибирск, Россия

²Институт вычислительных технологий СО РАН, Новосибирск, Россия

³Новосибирский государственный университет, Россия

*Контактный e-mail: rakitsky.anton@gmail.com

Предложен теоретико-информационный подход к оценке производительности суперкомпьютеров. Этот подход основан на характеристике “вычислительная способность”, которую можно определить теоретически, опираясь исключительно на описание архитектуры исследуемой системы. Предложен метод определения вычислительной способности суперкомпьютеров с учетом влияния межпроцессорного взаимодействия. Метод применяется в работе для оценки производительности пяти суперкомпьютеров среди первых 50 позиций в списке TOP500 (ноябрь 2016). Полученные результаты сравниваются с аналогичными значениями бенчмарка LINPACK и характеристиками Theoretical peak, используемыми для формирования рейтинга TOP500. По результатам сравнения делаются выводы относительно применимости указанного подхода для оценки производительности суперкомпьютеров на этапе разработки.

Ключевые слова: производительность, вычислительная способность, суперкомпьютер, теория информации, процессор.

Введение

Оценка производительности многопроцессорных вычислительных систем (суперкомпьютеров) является очень важной задачей как для проектировщиков подобных систем, так и для их пользователей. Существует множество бенчмарков [1, 2], которые позволяют дать оценку подобных систем относительно решения различных задач, но применение бенчмарков требует наличия рабочей модели исследуемой системы, поэтому их практически невозможно использовать на этапе разработки. Помимо бенчмарков существует и теоретический подход для оценки производительности — *теоретический максимум*. В этом случае используется характеристика IPC (Instructions Per Cycle), которая, как правило, предоставляется производителем. Характеристика IPC — это среднее количество инструкций, которое процессор способен выполнить за один такт, что является очень грубой и неточной оценкой производительности, к тому же эта характеристика может варьироваться в зависимости от выбранного метода ее определения.

В данной статье предложен теоретический подход к оценке производительности современных многопроцессорных вычислительных систем (суперкомпьютеров), который не требует проведения экспериментов над рабочей моделью исследуемой вычислитель-

ной системы и может стать хорошей альтернативой существующим методам. Теоретический базис метода впервые представлен в работе [3], позднее он был применен для оценки производительности реальных компьютеров на базе процессоров Intel [4]. Идея метода близка к основным положениям теории информации Шеннона [5], в качестве меры производительности она использует такую характеристику, как *вычислительная способность*.

В настоящей работе описывается применение представленного метода для оценки производительности современных суперкомпьютеров. Ранее метод уже применялся к суперкомпьютерам, однако рассматривались только модели с простейшей архитектурой, в которых сетевое взаимодействие учитывается относительно просто [6]. Кроме того, в статье уточняются особенности оценки вычислительной способности суперкомпьютеров, приводятся формальное описание и подробные формулы, описывается принцип построения уравнения сетевого процессора, приводится подробное описание построения уравнений для межпроцессорных соединений Aries interconnect и Infiniband FDR, оценивается вычислительная способность суперкомпьютера, содержащего в вычислительных узлах Intel Xeon Phi — многоядерные процессоры, имеющие до 72 вычислительных ядер, разработанные специально для параллельных вычислений. Описан процесс оценки вычислительной способности пяти суперкомпьютеров из списка TOP500 (ноябрь 2016 г.). Исследуются суперкомпьютеры из списка: Toraz [7] (27-е место), Shaheen II [8] (15-е место), Hazel Hen [9] (14-е место), Trinity [10] (10-е место) и Cori [11] (5-е место). Последние четыре модели из списка относятся к семейству Cray XC40, поэтому, чтобы показать надежность описанного метода, они сравниваются с суперкомпьютером Toraz. Toraz имеет ряд существенных отличий от остальных моделей, например, для межпроцессорного взаимодействия и обмена данными используется сеть Infiniband FDR, в то время как остальные суперкомпьютеры построены на базе Aries interconnect. Предварительные результаты по оценке некоторых суперкомпьютеров докладывались на конференции [12].

1. Вычислительная способность

Основной теоретический базис представлен в работе [3], поэтому в данном разделе приводятся только самые необходимые определения и теоремы, без которых понимание дальнейшей работы представляется затруднительным.

Пусть имеется компьютер с памятью M и набором инструкций I . Каждая инструкция $s \in I$ состоит из имени и значений ее операндов. Из этого следует, что одна и та же инструкция с различными значениями операндов будет включена в набор I как независимая. Например, будем считать, что MOV R0 R1 и MOV R0 R2 (где R0, R1 и R2 — регистры) — это различные инструкции и обе входят в I .

Примем $\tau(s)$, $s \in I$, за время выполнения инструкции s и $S = s_1, s_2, \dots, s_n$, $s_i \in I$, будем считать компьютерной задачей. Тогда время выполнения задачи S определяется как $\tau(S) = \sum_{i=1}^n \tau(s_i)$. Кроме того, мы считаем, что все $\tau(s)$ — это целые числа и их наибольший общий делитель равен 1 (такое допущение верно практически для всех реальных процессоров, так как в их списке инструкций присутствует инструкция, чье время выполнения составляет 1 единицу времени, т. е. $\tau(s) = 1$). Определим количество различных задач, чье время выполнения равно T , как $\nu(T)$ и отметим, что оно так же равно размеру множества всех последовательностей инструкций, чье время выполнения равно T , т. е. $\nu(T) = N(T)$, где $N(T) = |\{S = s_1, s_2, \dots, s_n : s_i \in I, \tau(S) = T\}|$.

Пусть есть процессор, который может выполнять N различных последовательностей инструкций с временем выполнения 1 мин. Мы можем сказать, что этот процессор способен выполнять N^2 последовательностей инструкций за 2 мин, потому что если S_1 и S_2 — это одноминутные последовательности, то объединенная последовательность S_1S_2 является двухминутной (мы не рассматриваем некоторые дополнительные двухминутные последовательности, у которых инструкция начинает выполняться в конце первой минуты, а заканчивается в начале второй). Аналогичным образом, порядка N^k последовательностей могут быть выполнены за k минут. Таким образом, количество возможных последовательностей растет экспоненциально как функция от времени T ($N(T) \approx 2^{CT}$), следовательно, $\log N(T)/T$ (или предел этого значения) — это адекватная мера мощности процессора и вычислительная способность определяется как

$$C(I) = \lim_{T \rightarrow \infty} \frac{\log N(T)}{T}. \quad (1)$$

Важным вопросом является, как определить или хотя бы оценить значение $C(I)$ из (1). Очевидно, что прямое вычисление предела невозможно, однако в комбинаторном анализе существует метод для определения $C(I)$. Шеннон использовал этот метод [5], когда оценивал пропускную способность канала. Рассмотрим набор инструкций I как алфавит и будем считать, что все слова из этого алфавита (последовательности символов) допустимы и могут быть выполнены на компьютере. Это допущение позволяет оценить верхнюю границу вычислительной способности, так как для любого процессора множество допустимых последовательностей инструкций является подмножеством всех возможных последовательностей. Учитывая все вышесказанное, $C(I)$ равна логарифму наибольшего действительного решения X_0 следующего характеристического уравнения:

$$X^{-\tau(s_1)} + X^{-\tau(s_2)} + \dots + X^{-\tau(s_n)} = 1. \quad (2)$$

В работе [3] показано, что вычислительная способность многоядерных процессоров определяется как сумма вычислительных способностей каждого ядра.

Суперкомпьютеры и подобные вычислительные системы можно рассматривать как набор вычислительных узлов, соединенных по сети. Ожидается, что все вычисления выполняются на подобных системах параллельно, а следовательно, каждый узел (более того, каждое вычислительное ядро в каждом узле) по сути выполняет все действия независимо, кроме взаимодействия с другими узлами.

Вычислительный узел состоит из процессоров, оперативной памяти и сетевого адаптера (в данной работе мы его будем называть сетевым процессором — СП), размещенных на одной плате. В работе [3] показано, что если есть вычислительная система с $N \geq 1$ вычислительными узлами I_1, \dots, I_N и каждый узел может работать самостоятельно и независимо от других, то вычислительная способность этой системы представляет собой сумму вычислительных способностей узлов I_1, \dots, I_N . Таким образом, вычислительная способность многопроцессорной системы определяется следующим образом:

$$C_{\text{в.с}} = \sum_{i=1}^N C_i,$$

где N — количество узлов, C_i — вычислительная способность i -го узла. В данном случае вычислительная способность узла определяется как

$$C_i = C_{\text{СП}_i} + \sum_{j=1}^{N_i} C_{\text{core}_j},$$

где N_i — количество ядер в i -м узле; C_{core_j} — вычислительная способность j -го ядра; $C_{\text{СП}_i}$ — вычислительная способность сетевого процессора, расположенного в i -м узле.

2. Вычислительная способность современных процессоров

Рассмотрим две важные особенности современных процессоров, которые необходимо учитывать при построении уравнения (2), — это кэш-память и конвейер.

2.1. Кэш-память

Современные процессоры включают в свое ядро память с быстрым доступом. Эта память, называемая кэш-памятью, используется, чтобы снизить среднее время доступа к данным из основной памяти. Когда процессор обращается к какому-то участку памяти, в первую очередь проверяется, не был ли этот участок загружен в кэш-память, и только если его там нет, запрос отправляется уже к основной памяти. Запрашиваемые участки памяти хранятся в кэш-памяти согласно алгоритму кэширования (он может различаться для разных процессоров).

Рассмотрим компьютер с набором инструкций I , основной памятью M и кэш-памятью L . Пусть есть инструкция $s \in I$ с одним операндом — клеткой памяти, тогда $\tau(s)$ — базовое время выполнения инструкции (без обращения к памяти), $\tau(M)$ — время обращения к основной памяти и $\tau(L)$ — время обращения к кэш-памяти. В таком случае инструкция s будет представлена в характеристическом уравнении (2) следующим образом:

$$\dots + \frac{|L|}{X^{\tau(s)+\tau(L)}} + \frac{|M| - |L|}{X^{\tau(s)+\tau(L)+\tau(M)}} + \dots \quad (3)$$

Здесь $|L|, |M|$ — количество различных ячеек кэш- и основной памяти, рассчитанное относительно размера операнда инструкции. Поясним представленные в уравнении (3) слагаемые. При выполнении инструкции возможны две ситуации, когда искомая ячейка находится в кэш-памяти и когда ее там нет. Первая ситуация описывается первым слагаемым. Если в инструкции описана ячейка памяти, которая находится в кэше, то количество таких ячеек равно $|L|$, а время выполнения инструкции будет равно базовому времени плюс время обращения к кэш-памяти. Во втором же случае количество различных ячеек составит $|M| - |L|$, а время выполнения — $\tau(s) + \tau(L) + \tau(M)$, так как процессор в любом случае обратится сначала к кэш-памяти и только потом к основной. Аналогичным образом можно расширить описание и для многоуровневой кэш-памяти. Примеры и более подробный анализ влияния кэш-памяти с несколькими уровнями на вычислительную способность представлены в статье [4].

2.2. Конвейер

Все современные процессоры строятся на базе конвейера, основной задачей которого является уменьшение времени выполнения инструкций. Достигается это за счет выполнения инструкции по частям. Конвейер представляет собой последовательность блоков

обработки данных, в которой выходные данные одного блока являются входными данными следующего. Каждый блок работает независимо от других. Например, если одна инструкция обрабатывается в блоке декодирования, то другая может обрабатываться в блоке выполнения. Обработка инструкции в каждом конкретном блоке происходит существенно быстрее, чем выполнение всей инструкции целиком, поэтому, несмотря на то, что время выполнения одной единственной инструкции возрастает, время выполнения последовательности инструкций (процессорной задачи) существенно уменьшается. Для нашего метода важно правильно определять время выполнения инструкции. И с учетом наличия в процессоре конвейера время выполнения инструкции мы будем определять как максимальную задержку конвейера, которая может возникнуть из-за этой инструкции (очевидно, что пока одна инструкция выполняется в блоке, следующая за ней вынуждена ждать, пока блок освободится).

В современных процессорах Intel (а именно на их базе построены рассматриваемые в данной работе суперкомпьютеры) инструкции в процессе декодирования разбиваются на простейшие микрооперации, каждая из которых на каждом блоке выполняется за один такт, за исключением операций обращения к памяти, к которым необходимо прибавить время обращения к памяти. Поэтому временем выполнения инструкции в большинстве случаев является количество микроопераций, на которые она разбивается. Подробное описание влияния длины конвейера, разбиения инструкций на микрооперации и прочих особенностей конвейеров современных процессоров с приведением подробных примеров представлено в работе [4].

3. Процессоры Haswell

Чтобы оценить вычислительную способность процессоров на базе микроархитектуры Haswell, необходимо рассмотреть их основные особенности. Процессоры Haswell содержат регистровый файл, в котором 168 целочисленных регистров и 168 векторных (они же используются и для вещественных операций). Это означает, что инструкция имеет 168 целочисленных регистров в качестве значения операнда вместо 16 регистров (как было в ранних процессорах). Другой важной особенностью является пропускная способность блоков конвейера Haswell, она составляет четыре микрооперации за один такт. Следовательно, внутри процессора могут одновременно выполняться до четырех цепочек из связанных микроопераций. В этих цепочках следующая микрооперация зависит от результатов работы предыдущей и не может быть обработана раньше ее. В нашем методе оценивается верхняя граница вычислительной способности, поэтому можно считать, что в процессоре четыре независимых конвейера, и это означает, что после определения вычислительной способности процессора (уравнение строится для одного конвейера) полученное значение необходимо умножить на 4.

Опишем процесс построения характеристического уравнения (2) для процессоров Haswell. В первую очередь формируется файл со списком инструкций процессора в базовом формате. В этом формате мы описываем инструкцию как имя, список типов операторов и базовое время ее выполнения. Полученный файл содержит 681 запись. Затем при помощи программы преобразуем этот файл в список инструкций в итоговом формате. В этом формате в инструкцию уже подставляются числа на место операндов. Например, инструкция “MOV r,r 1” преобразуется в “MOV 28224 1”. Число 28224 — это количество различных инструкций (т.е. различных комбинаций значений операндов) “MOV r,r 1” из набора инструкций I , так как каждый операнд “r” может принимать

одно из 168 различных значений, следовательно, количество различных комбинаций равно $168 \cdot 168 = 28\,224$.

Для автоматического преобразования файла из базового формата в итоговый используется программа на основе алгоритма рекурсивного спуска. Этот алгоритм применяется для корректного разбора сложных инструкций, например, инструкция `MOV r8/r16,r8/r16` представляет собой набор инструкций: $\{\text{MOV r8,r8}; \text{MOV r8,r16}; \text{MOV r16,r8}; \text{MOV r16,r16}\}$. Файлы со списками инструкций в базовом и итоговом форматах, а также программу для их преобразования можно найти по адресу [13].

После формирования необходимых файлов применяется программа оценки вычислительной способности ядра Haswell с одним конвейером $C_1(I) \approx \log_2 524\,014\,684.672 \approx 28.965$ бит/такт. Поскольку каждое ядро имеет четыре конвейера, $C = C_1 \cdot 4 \approx 28.965 \cdot 4 \approx 115.86$ бит/такт. Чтобы оценить вычислительную способность конкретного процессора, необходимо полученное значение умножить на его тактовую частоту, чтобы преобразовать единицы измерения в бит/с. Так как процессоры Haswell многоядерные, полученное значение необходимо умножить на количество ядер конкретного процессора.

4. Сетевой процессор

Основная задача сетевого процессора — формирование пакета данных и отправка его по сети другому сетевому процессору или, наоборот, получение пакета данных от некоторого сетевого процессора. Таким образом, набор инструкций СП будет состоять из инструкций, которые отсылают некоторый пакет данных от одного узла другому (как правило, вычислительный узел содержит только один сетевой процессор) или получают их. С учетом описанных особенностей сетевого процессора получим характеристическое уравнение для СП, находящегося в вершине с номером k :

$$\sum_{i=\text{minSize}}^{\text{maxSize}} \sum_{j=1, j \neq k}^N \frac{M_{i,k} + M_{i,j}}{X^{T_{i,j}}} = 1, \quad (4)$$

где minSize и maxSize — минимальный и максимальный возможные размеры пакета в сети; N — количество узлов суперкомпьютера; $M_{i,j}$ — количество всевозможных различных пакетов размера i , которые могут быть сформированы в узле j ; T_j — время передачи пакета размером i между узлом k и узлом j . Числитель $M_{i,k} + M_{i,j}$ сформирован таким образом, потому что включает как инструкции отправки ($M_{i,k}$), так и инструкции приема ($M_{i,j}$).

4.1. Сеть Aries interconnect

Четыре из исследуемых суперкомпьютеров относятся к семейству Cray XC40, они используют сеть Aries interconnect для межпроцессорного взаимодействия. Рассмотрим структуру этой сети. Система состоит из нескольких стоек, каждая из них содержит в себе три блэйд-шасси по 16 блэйдов в каждом, блэйд в свою очередь состоит из четырех двухпроцессорных узлов, каждый узел содержит до двух вычислительных процессоров, общую оперативную память и сетевой процессор. Сеть формируется из электрических групп по две стойки (384 узла в каждой). Все соединения внутри группы электрические и позволяют передавать данные со скоростью 14 Гбит/с по каждой линии. Все группы

соединены 12-жильными оптическими кабелями, обеспечивающими скорость передачи 12.5 Гбит/с. Подробное описание технических характеристик можно найти в [14]. Рассмотрим наиболее важные факторы, которые влияют на уравнение (4):

- Сетевой процессор разбивает все запросы на пакеты и отправляет эти пакеты по сети, каждый пакет содержит не более 64 байт данных.
- Скорость передачи данных по электронному соединению составляет 14 Гбит/с.
- Скорость передачи данных по оптоволоконному соединению 12.5 Гбит/с.
- Взаимодействие между процессорами обеспечивается при помощи сетевых процессоров (кроме тех, которые расположены в одном узле).

Опираясь на эти характеристики, можно построить уравнение (4), которое будет идентичным для всех СП суперкомпьютера:

$$\sum_{i=1}^{64} \left(\frac{3M_i}{X^{2T_{el_i}}} + \frac{(N_{gr} - 4)M_i}{X^{3T_{el_i}}} + \frac{(N_{nd} - N_{gr})M_i}{X^{2T_{el_i} + T_{op_i}}} \right) = 1. \quad (5)$$

Здесь i — размер пакета в байтах; M_i — количество различных пакетов размера i , которые можно сформировать в одном узле; T_{el_i} — время передачи пакета данных размером i байт по электронному соединению; T_{op_i} — время передачи пакета данных размером i байт по оптоволоконному соединению; N_{gr} — количество узлов в группе; N_{nd} — общее число узлов суперкомпьютера.

В уравнении (5) можно выделить три варианта передачи данных между узлами.

1. Данные передаются между двумя узлами внутри одного блэйда. В данном случае будет всего две пересылки: от источника к роутеру и от роутера к узлу назначения. Пересылки осуществляются по электронным соединениям.
2. Данные передаются между узлами внутри одной группы, но из разных блэйдов. В таком случае будет три передачи: от источника к его роутеру, от роутера источника к роутеру узла назначения и от роутера к узлу назначения. Все передачи будут происходить по электронным соединениям.
3. Данные передаются между узлами из разных групп. Как и в предыдущем пункте, будут три передачи: от источника к его роутеру, от роутера источника к роутеру узла назначения (так как они в разных группах, передача будет идти по оптоволоконному соединению) и от роутера к узлу назначения.

Таким образом, случай 1 представлен в уравнении (5) слагаемым $\frac{3M_i}{X^{2T_{el_i}}}$ (здесь 3 — количество узлов на блэйде без учета текущего), случай 2 представлен слагаемым $\frac{(N_{gr} - 4)M_i}{X^{3T_{el_i}}}$ и случай 3 — слагаемым $\frac{(N_{nd} - N_{gr})M_i}{X^{2T_{el_i} + T_{op_i}}}$.

4.2. Сеть InfiniBand FDR

Суперкомпьютер Toraz-SGI ICE X построен на основе сети InfiniBand FDR [15]. Метод оценки вычислительной способности сетевого процессора для InfiniBand аналогичен описанному ранее. Рассмотрим важные для уравнения (4) технические характеристики:

- минимально возможный размер пакета данных для передачи составляет 32 байта [15];
- максимально возможный размер пакета данных для передачи 4096 байт;
- теоретическая пропускная способность одиночного соединения 13.64 Гбит/с.

Вычислительная способность СП для сети InfiniBand FDR у исследуемого суперкомпьютера Toraz составила порядка 9.24 Гбит/с.

4.3. Суперкомпьютер Cori

Главным отличием суперкомпьютера Cori является наличие двух модулей, один из которых построен полностью на процессорах Intel Xeon Phi (Knights Landing), а второй — на процессорах Haswell, т. е. в суперкомпьютере присутствуют узлы разных типов, что необходимо учитывать в итоговом уравнении. Процессор Intel Xeon Phi — это многоядерный сопроцессор, разработанный Intel специально для многопроцессорных вычислительных систем. В рассматриваемой системе используется 68-ядерная версия процессора. Конвейер ядра построен на основе микроархитектуры Silvermont с небольшими доработками и улучшениями, однако по производительности ядро этого процессора существенно уступает процессору Haswell. Вычислительная способность первого потока этого процессора равна 22.4925 бит/с, тем не менее он может поддерживать до четырех одновременных потоков в каждом ядре, поэтому необходимо указанное значение умножить на 4. Огромным преимуществом этих процессоров по отношению к параллельным вычислениям является поддержка векторных инструкций, которые позволяют оперировать с данными объемом до 512 бит.

5. Анализ результатов

Чтобы показать эффективность характеристики “вычислительная способность”, мы исследовали пять суперкомпьютеров: Toraz, Shaheen II, Hazel Hen, Trinity и Cori. Важно отметить, что это суперкомпьютеры трех разных типов: Toraz соответствует первому типу (соединение InfiniBand FDR), Shaheen II, Hazel Hen и Trinity — типичные представители семейства суперкомпьютеров Cray XC40, а суперкомпьютер Cori является представителем суперкомпьютеров с узлами различной конфигурации. Характеристики и результаты исследованных суперкомпьютеров представлены в табл. 1, где C — вычислительная способность одного сетевого процессора. В столбце для суперкомпьютера Cori встречаются по два числа, первое относится к модулю на процессорах Haswell, а второе — к модулю на процессорах Xeon Phi. LINPACK — название бенчмарка, представленного в списке TOP500 как Rmax. Theoretical peak — характеристика, называемая Rpeak в списке TOP500.

Т а б л и ц а 1. Характеристики суперкомпьютеров

Показатель	Toraz	Shaheen 2	Hazel Han	Trinity	Cori
Количество ядер	124200	196608	185088	301056	73728; 678912
Тактовая частота, МГц	2600	2300	2500	2300	2300; 1400
C , Мбит/с	9239.28	1063541.62	1064558.14	1065457.03	1066894.66; 1066750.81
LINPACK, Тфлопс/с	3318.95	5536.99	5640.17	8100.90	14014.70
Theoretical peak, Тфлопс/с	4570.56	7235.17	7403.52	11078.90	27880.70
Вычислительная способность, Гбит/с	37445.43	58926.16	61820.67	90248.71	112944.79

Т а б л и ц а 2. Относительные результаты характеристик

Показатель	Тораз	Shaheen II	Hazel Han	Trinity	Cori
Вычислительная способность	1.00	1.574	1.651	2.410	3.016
LINPACK	1.00	1.668	1.699	2.441	4.223
Theoretical peak	1.00	1.583	1.620	2.424	6.100

Единицы измерения этих характеристик различны, поэтому невозможно напрямую сравнить вычислительную способность LINPACK и Theoretical peak. Во-первых, упорядочим суперкомпьютеры по возрастанию вычислительной способности. Следующим шагом каждое значение разделим на соответствующее значение суперкомпьютера Тораз. Таким образом, получим табл. 2, в которой все значения характеристик, соответствующих суперкомпьютеру Тораз, будут равны 1. Представленные в табл. 2 значения относительны и не имеют единиц измерения, поэтому их можно легко сравнивать друг с другом.

Более наглядно результаты представлены на диаграмме (см. рисунок). Из диаграммы следует, что полученные результаты вычислительной способности ближе к значениям бенчмарка, чем Theoretical peak. Важно отметить, что значение бенчмарка LINPACK является основной характеристикой в рейтинге TOP500 и места суперкомпьютеров распределяются исходя из этого значения (Theoretical peak используется только в случае равенства значений бенчмарка).

На диаграмме можно заметить, что для суперкомпьютера Cori вычислительная способность показала меньший рост, чем остальные характеристики. Обуславливается это в первую очередь узкой специализацией процессоров Intel Xeon Phi, которые позволяют существенно ускорить легко распараллеливающиеся задачи, однако с остальными задачами они справляются гораздо слабее обычных процессоров. Из-за особенностей архитектуры и списка инструкций такие процессоры показывают очень хорошие результаты для задачи решения систем линейных алгебраических уравнений (способность решать эту задачу и исследуется в рамках бенчмарка LINPACK), тогда как на других задачах, которые сложнее распараллелить, дают более слабый результат.

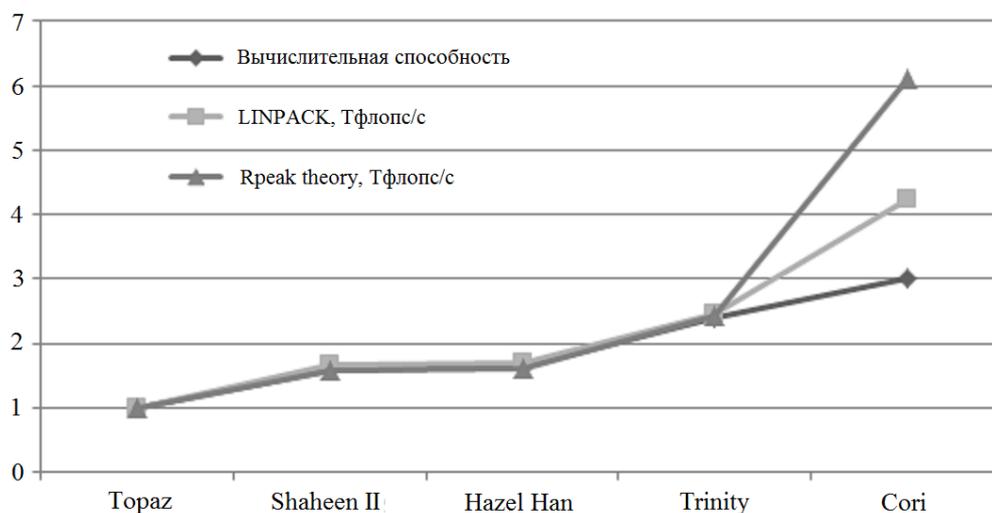


Диаграмма данных по табл. 2

Так как предложенная в работе характеристика — вычислительная способность — оценивает способность компьютера решать все возможные виды задач, то, очевидно, что результаты такого показателя будут ниже, чем результаты применения бенчмарка LINPACK, сама суть которого сводится к очень хорошо распараллеливаемой задаче решения систем линейных уравнений.

Таким образом, влияние процессоров Intel Xeon Phi, находящихся в узлах Cori, на вычислительную способность всего суперкомпьютера будет меньше, чем на результаты применения бенчмарка LINPACK. При этом разница между значениями вычислительной способности и бенчмарка LINPACK для этого суперкомпьютера все равно существенно меньше разницы между LINPACK и Theoretical peak.

Полученные результаты позволяют утверждать, что метод теоретической оценки вычислительной способности современных многопроцессорных вычислительных систем (без проведения каких-либо экспериментов над рабочей моделью исследуемой системы) является достаточно точным и может быть использован для оценки их производительности. Кроме того, он может быть использован и при разработке новых вычислительных систем на стадии проектирования, когда подбираются оптимальные компоненты и способ соединения. Данный метод позволяет легко варьировать параметры и быстро пересчитывать значение вычислительной способности.

В настоящее время авторами исследуются такие направления, как возможность применения предложенного метода к графическим процессорам, процессорам с MIPS-архитектурой и современным ARM-процессорам, что позволит объединить большинство существующих вычислительных процессоров при помощи единой теоретической методики оценки производительности.

Благодарности. Работа выполнена при финансовой поддержке РФФИ (грант № 15-07-0185), Президентской программы “Ведущие научные школы РФ” (грант № НШ-7214.2016.9) и правительства Новосибирской области.

Список литературы / References

- [1] **Weicker, R.P.** A detailed look at some popular benchmarks // *Parallel Computing*. 1991. Vol. 17. P. 1153–1172.
- [2] **Lilja, D.J.** *Measuring computer performance: A practitioner’s guide*. Cambridge Univ. Press, 2005. 280 p.
- [3] **Ryabko, B.** An information-theoretic approach to estimate the capacity of processing units // *Performance Evaluation*. 2012. Vol. 69. P. 267–273.
- [4] **Ryabko, B., Rakitskiy, A.** An analytic method for estimating the computation capacity of computing devices // *J. of Circuits, Syst. and Comput.* (JCSC). Vol. 26, No. 5. 2017. P. 1750086. Doi:10.1142/S0218126617500864.
- [5] **Shannon, C.E.** A mathematical theory of communication // *Bell Syst. Techn. J.* 1948. Vol. 27. P. 379–423.
- [6] **Ракитский А.А., Рябко Б.Я., Фионов А.Н.** Аналитический метод сравнения и оценки производительности вычислительных систем // *Вычисл. технологии*. 2014. Т. 19, № 4. С. 84–98.
Rakitskiy, A.A., Ryabko, B.Ya., Fionov, A.N. The analytical method for comparing and evaluating the performance of computers and computer systems // *Comput. Technologies*. 2014. Vol. 19, No. 4. P. 84–98.
- [7] Description of Topaz supercomputer. Available at: <http://top500.org/system/178523>

- [8] Description of Shaheen II supercomputer. Available at: <http://top500.org/system/178515>
- [9] Description of Hazel Hen supercomputer. Available at: <http://top500.org/system/178446>
- [10] Description of Trinity supercomputer. Available at: <http://top500.org/system/178610>
- [11] Description of Cori supercomputer. Available at: <http://top500.org/system/178924>
- [12] **Rakitskiy, A., Ryabko, B.** An information-theoretic approach to performance evaluation of supercomputers // XV Intern. Symp. "Problems of Redundancy in Inform. and Control Syst." (REDUNDANCY), St. Petersburg, 2016. P. 125–128. Doi:10.1109/RED.2016.7779345.
- [13] Laboratory of information systems and data protection. Files with equations for processors. Available at: <http://www.ict.nsc.ru/ru/structure/orgunits/lab-info-sys-security-page>
- [14] **Alverson, B., Froese, E., Kaplan, L., Roweth, D.** Cray XC Series Network. Available at: www.cray.com/sites/default/files/resources/CrayXCNetwork.pdf
- [15] **Shanley, T.** InfiniBand Network architecture. Addison-Wesley, 2003. 1208 p.

*Поступила в редакцию 27 июня 2017 г.,
с доработки — 13 ноября 2017 г.*

Informational-theoretical approach to the evaluation of supercomputers

RAKITSKIY, ANTON A.^{1,2,*}, RYABKO, BORIS YA.^{2,3}

¹Siberian State University of Telecommunications and Information Sciences,
Novosibirsk, 630102, Russia

²Institute of Computational Technologies SB RAS, Novosibirsk, 630090, Russia

³Novosibirsk State University, Novosibirsk, 630090, Russia

*Corresponding author: Rakitskiy, Anton A., e-mail: rakitsky.anton@gmail.com

In this paper we present the informational-theoretical approach to the evaluation of performance for supercomputers. This approach is based on the computer capacity characteristic which can be estimated theoretically (without any experiments over the working model of the investigated computer) relying on the description of the architecture of the system under study. We describe the method for estimation of the supercomputers Computer capacity which takes into account the influence of the interprocess communication. The described method is used in this paper to estimate the performance of the five supercomputers among the first fifty positions in TOP500 list (according to November 2016 data).

The obtained results are compared with the values of LINPACK benchmark and Theoretical peak characteristic which are used to form the TOP500 rating. Based on the comparison we draw conclusions about the applicability of the presented approach to the evaluation of supercomputers performance at the design stage.

Keywords: performance, computer capacity, supercomputer, information theory, processor.

Acknowledgements. This research was partly supported by RFBR (grant № 15-07-0185), the Ministry of Education and Science of RF (grant No. NSH-7214.2016.9) and by the Government of the Novosibirsk Region.

Received 27 June 2017

Received in revised form 13 November 2017