

НРС — инструмент мирового развития

А. ГОЛЫШКО, канд. техн. наук, г. Москва

"Существует и лучший мир, но там всё очень дорого".

(Из учебника по "Цифровой экономике")

Термин НРС (High-Performance Computing — высокопроизводительные вычисления) появился сравнительно недавно и представляет собой раздел прикладной информатики, занимающийся поиском путей решения задач, требующих значительных вычислительных ресурсов. Прежде всего, там, где необходимо решать величайшие мировые научные проблемы, и инженерные и учёные умы из каждой области знаний могут подсказать какие именно. Очень сложные задачи всё ещё остаются нерешёнными без серьёзных вычислительных ресурсов. К ним относятся прогнозирование природных катастроф, лечение пока ещё неизлечимых заболеваний, исследования вселенной и многое другое. В настоящее время растёт интерес к НРС со стороны бизнеса, и уже появляются успешные примеры использования данной технологии.

В развитых странах на самом высоком уровне НРС уже используются в интересах государственной безопасности и реагирования на различные угрозы от техногенных аварий и прочих ЧС до терроризма, а также прогнозирования социальных, политических и прочих последствий тех или иных экономических преобразований. На следующем уровне — промышленная и исследовательская сферы, которые включают в себя промышленное производство, здравоохранение, нефтегазодобычу, естественные науки и, разумеется, финансы.

Развитие НРС долгое время шло в контексте очень ограниченного числа сценариев их практического использования, что было связано, в том числе, и с дефицитом соответствующей вычислительной техники. Достигнутый технический прогресс налицо, и многие востребованные в самых разных отраслях инновационные разработки создаются с помощью НРС. Вместе с этим меняется роль ИТ-компаний, которые превратились из довольно немногочисленного "клуба любителей суперкомпьютеров" и решения отдельных научных проблем в поставщиков, весьма востребованных в современном цифровом мире услуг. Поэтому сфера НРС развивается невероятно высокими темпами, и многие заказчики решений НРС, активно использующие локальные центры обработки данных, начинают внедрять

облачные вычисления, способные повысить эффективность работы конечных пользователей, гибкость и экономичность.

Предприятия многих отраслей совсем недавно получили практическую возможность собирать и хранить огромный объём первичных данных, детально описывающих буквально каждое движение, которое кем-либо и для каких-либо целей предпринимается в бизнесе. Возможность обработки этих данных с целью получения не только понятного для бизнеса, но и при этом на 100 % обоснованного результата (в большинстве случаев либо классификация текущих событий, либо прогноз на будущее) — довольно сложная задача, которая по большей части ещё только обсуждается в профессиональных кругах. Впрочем, с появлением практических результатов интерес к данной теме возрастает. В настоящее время появляется всё больше специалистов в данной области, которые на основе имеющихся у них академических знаний создают математические модели, применимые для решения той или иной бизнес-задачи. Эти решения "обкатываются" в отраслевых сообществах и на индустриальных площадках, проверяются, уточняются, классифицируются. Сама дисциплина Machine Learning (машинное обучение, являющееся частью разработки искусственного интеллекта) делится на отдельные поднаправления, среди которых одним из наиболее заметных является Deep Learning (глубокое обучение), и идеология НРС, разумеется, находит здесь своё место. Так постепенно создаётся весьма зрелая инфраструктура обработки первичных данных для целей бизнеса и формируется своего рода промышленный конвейер движения этих данных к конечному бизнес-результату.

Считается, что окупаемость НРС более чем в два раза превышает вложенные капитальные средства. В частности, на производстве объединение и дополнение физических моделей САПР с цифровыми двойниками данными с датчиков Интернета вещей и моделями искусственного интеллекта и глубокого обучения позволяет выявлять и прогнозировать аномальное поведение продуктов и их сбой. Эта функция профилактического обслуживания помогает выполнять оптимальные корректирующе-

щие действия, повышать качество технического обслуживания продуктов и ускорять разработку новых продуктов. В нефтегазовой отрасли, требующей высокопроизводительных вычислений, рабочие задачи, такие как обработка сейсмических данных и имитация функционирования резервуаров, интегрируются со средствами аналитики, что позволяет оптимизировать добычу и поставку сырой нефти и обеспечивает снижение стоимости одного барреля продукта.

Объединение инструментов 3D-моделирования и симуляции, а также интеграция и анализ информации, взятой из историй болезней пациентов и многих других пулов данных (геномных данных, данных визуализации, клинических, поведенческих данных, данных об эффективности лекарственных средств, медицинских отчётов и т. д.), намного повышают качество лечения. Более того, такой подход позволяет составлять планы лечения, проводить хирургические процедуры, совершенствовать конструкцию медицинского оборудования, получать более высокие результаты лечения пациентов, а также помогает переводить отрасль здравоохранения на модель медицинского обслуживания с высокой ценовой эффективностью.

Для успешной рыночной деятельности требуется проводить множество сложных операций финансового моделирования, включающих тысячи вычислительных задач длительностью от нескольких секунд до нескольких минут. Обычно эти вычисления выполняются в крупномасштабных средах НРС и требуют сотни мегабайт данных, таких как историческая стоимость обыкновенных акций за несколько лет. Однако для каждой задачи обычно используется только небольшая часть этих данных, которая должна обязательно интегрироваться со средствами анализа рисков, чтобы финансовые организации могли успешно соблюдать строгие и сложные нормативные требования.

Тем временем драйверами дальнейшего роста использования НРС выступают задачи моделирования работы человеческого мозга и точного прогноза погоды. Требуемая для них мощность превосходит сегодняшние ресурсы в 100000 раз, и её достижение прогнозируется лишь к 2030 г.

В целом НРС используют сконцентрированную вычислительную мощность для решения задач, требующих большого объёма вычислений и данных, включая моделирование и рендеринг (процесс получения изображения по модели с помощью компьютерной программы), которые не по силам стандартным рабочим станциям. Иначе говоря, возникшие проблемы не могут быть решены на обычном компьютере в течение разумного периода времени или выполнение невозможно из-за ограниченности доступных ресурсов (требуется слишком много данных). Так вот НРС — это подход к преодолению подобных ограничений путём использования специализированного или высокопроизводительного оборудования (суперкомпьютеров) или накопления



вычислительной мощности от нескольких устройств.

Как известно, многие сложные задачи можно решать методом параллельных вычислений, когда любой отдельно взятый процесс выполняется независимо от результатов выполнения других процессов. А для этого можно воспользоваться несколькими относительно небольшими вычислительными системами (узлами), сгруппированными в вычислительные кластеры. Вычислительный кластер — это группа компьютеров (серверов), которые работают вместе и составляют единый унифицированный вычислительный ресурс. Каждый узел работает под управлением своей копии операционной системы. Такие вычислительные кластеры позволяют быстро и с относительно небольшими затратами решать комплексные задачи и могут располагаться целиком в центре обработки данных (ЦОД). Подобный подход широко применяется в финансовом моделировании, проектировании и компьютерной анимации. Подчас такие термины, как HPC-система, высоконагруженная система, высоконагруженный кластер, highload-система и суперкомпьютер обычно используются в рассматриваемой здесь области знаний как синонимы.

Благодаря увеличению числа центральных процессоров (ЦП) и узлов увеличивается вычислительная мощность, которая позволяет за единицу времени выполнять больше операций. Эффект увеличения вычислительной мощности называется ускорением, определяемым как соотношение между временем выполнения в параллельной системе и временем выполнения в последовательной системе.

При построении кластера встаёт задача, как распределять нагрузку между серверами. Для этого используется балансировка нагрузки, которая, кроме самого распределения, выполняет ещё ряд других задач, например, повышение отказоустойчивости (при выходе из строя одного из серверов система будет продолжать работать) и защита от некоторых типов атак.

В машинах с общей памятью и произвольным доступом (RAM) память может быть доступна всем процессорам. Между тем в кластерах распределённой памяти она недоступна между различными процессорами или узлами. При использовании распределённой памяти должно быть надёжное и скоростное сетевое соединение с малыми задержками для отправки сообщений между процессорами (или для использования других механизмов связи), поскольку они не имеют доступа к одному и тому же пространству памяти. Поэтому большинство кластеров HPC оснащены именно такими сетями, чтобы оперативно доставлять данные между различными узлами. Современные системы HPC часто представляют собой гибридную реализацию обеих концепций, поскольку некоторые устройства совместно используют общее пространство памяти, а некоторые нет.

Задачи, решаемые HPC, подразделяются на две категории: высокопроизводительные ("обычная" задача за

малое время или "тяжёлые" задачи, которые иначе решить невозможно) или большое количество "небольших" задач. Соответствующее распределение данных и операций между несколькими подразделениями требует концепции распараллеливания. Многие типы задач можно разделить на более мелкие части или подзадачи. Меньшие части могут быть построены, например, декомпозицией данных (параллелизм данных) или декомпозицией задач (параллелизм задач). Степень связи между этими подзадачами влияет на степень их распараллеливания.

Важным качественным показателем кластера является масштабируемость. Система является масштабируемой, если она способна увеличивать производительность пропорционально доступным ресурсам. Масштабируемость можно оценить через отношение прироста производительности системы к приросту используемых ресурсов. Чем ближе это отношение к единице, тем лучше. Также под масштабируемостью понимается возможность наращивания дополнительных ресурсов без структурных изменений центрального узла системы.

Масштабирование архитектуры высоконагруженной системы может быть горизонтальным и вертикальным. Вертикальное масштабирование заключается в увеличении производительности системы за счёт увеличения мощности сервера. Главный недостаток вертикального масштабирования в том, что оно ограничено определённым пределом. Ведь технические характеристики "железа" нельзя увеличивать бесконечно. Однако на самом деле вертикальная компонента присутствует практически всегда, а универсального горизонтального масштабирования как такового не существует. Горизонтальное масштабирование заключается в увеличении производительности системы за счёт подключения дополнительных серверов. Именно горизонтальное масштабирование является сейчас фактически стандартом. Существует также такой термин, как диагональное масштабирование, подразумевающее одновременное использование обоих подходов.

Современному заказчику нужно не просто уметь решать задачи, а решать их быстрее и лучше, чем это получалось на HPC предыдущего поколения. Поэтому хочется всё самое-самое от процессоров до сетей. Технические требования к HPC также находятся в постоянной динамике. Сегодня идёт борьба за экзафлопные вычисления, и уже пошли разговоры о зеттафлопных. А ведь, помимо "супержелеза", необходимо иметь соответствующее ПО, позволяющее максимально использовать вычислительные ресурсы. Помимо денег, на HPC нужны рабочие площади и электроэнергия. Тем более, что иным суперкомпьютерам уже требуются многие мегаватты, как ЦОДам. Если говорить по-хозяйски, то физически HPC-система должна помещаться в отведённом для неё зале и потреблять разумное количество электроэнергии.

Кстати, о флопах. При измерении производительности кластеров и су-

перкомпьютеров обычно используется пиковая производительность — теоретический предел производительности для данных процессоров и максимальная производительность, которую данный кластер или суперкомпьютер достигает при решении практических задач. FLOPS (флопс — производная от Floating-point Operations Per Second) — внесистемная единица, используемая для измерения производительности, показывающая, сколько операций с плавающей точкой в секунду выполняет данная вычислительная система. А ещё Flops — это общий тест для определения рейтинга скорости работы процессоров. Операции с плавающей точкой подразумевают любые операции, которые связаны с дробными числами. Такие операции занимают гораздо больше времени в отличие от операций с целыми числами. Большинство современных процессоров включают блок для вычисления операций с плавающей точкой (FPU), который является специализированной частью микропроцессора и отвечает за выполнение операций с плавающей точкой. Flops-измерения фактически измеряют скорость FPU.

Одним из важнейших достоинств показателя флопс является то, что он до некоторых пределов может быть истолкован как абсолютная величина и вычислен теоретически, в то время как большинство других популярных мер являются относительными и позволяют оценить испытываемую систему лишь в сравнении с рядом других. Эта особенность даёт возможность использовать для оценки результаты работы различных алгоритмов, а также оценить производительность вычислительных систем, которые ещё не существуют или находятся в разработке.

Несмотря на кажущуюся однозначность, в реальности флопс является достаточно плохой мерой производительности, поскольку неоднозначным является уже само его определение. Под операцией с плавающей запятой может скрываться масса разных понятий, не говоря уже о том, что существенную роль в данных вычислениях играет разрядность операндов, которая также нигде не оговаривается. Кроме того, величина флопс подвержена влиянию очень многих факторов, напрямую не связанных с производительностью вычислительного модуля, таких как пропускная способность каналов связи с окружением процессора, производительность основной памяти и синхронность работы кэш-памяти разных уровней.

Всё это, в конечном итоге, приводит к тому, что результаты, полученные на одном и том же компьютере с помощью разных программ, могут существенным образом отличаться. Более того, с каждым новым испытанием разные результаты можно получить при использовании одного алгоритма. Отчасти эта проблема решается соглашением об использовании единообразных тестовых программ. Самый распространённый тест для измерения производительности называется Linpack с усреднением результатов, но со временем возможности компьютеров перерастают рамки

принятого теста, и он начинает давать искусственно заниженные результаты, поскольку не задействует новейшие возможности вычислительных устройств. К некоторым системам общепринятые тесты вообще не могут быть применены, в результате чего вопрос об их производительности остаётся открытым. Кстати, Linpack используется при составлении рейтинга суперкомпьютеров TOP-500. Сам рейтинг появился сравнительно недавно, а весь перечень всемирных достижений в суперкомпьютерах приведён в **таблице**.

Некоторые ранние достижения в суперкомпьютерах выглядели так:

- ЭНИАК (1946 г.) — 300 флопс (при массе 27 т и энергопотреблении 150 кВт);
- IBM 709 (1957 г.) — 5 кфлопс;
- БЭСМ-6 (1968 г.) — 1 Мфлопс (операции деления);
- Cray-1 (1974 г.) — 160 Мфлопс;
- БЭСМ-6 на базе Эльбрус-1К2 (1980-е годы) — 6 Мфлопс (операции деления);
- Эльбрус-2 (1984 г.) — 125 Мфлопс;
- Cray-2 (1985 г.) — 1,9 Гфлопс;
- Cray Y-MP (1988 г.) — 2,3 Гфлопс;
- Электроника СС БИС (1991 г., двухмашинный вариант) — 500 Мфлопс;
- ASCII Red (1993 г.) — 1 Тфлопс;
- Fujitsu FX-1 (2008 г.) — 121 Тфлопс;
- VM Blue Gene/L (2006 г.) — 478,2 Тфлопс;
- Cray Jaguar (2008 г.) — 1,059 Пфлопс;
- IBM Roadrunner (2008 г.) — 1,042 Пфлопс;
- Ломоносов (2011 г.) — 1,3 Пфлопс; ну и так далее...

По состоянию на июнь 2019 г. верхняя часть TOP-500 остаётся в основном неизменной. Это — два суперкомпьютера, построенных IBM, Summit и Sierra, установленных в Национальной лаборатории Ок-Риджа (ORNL) Министерства энергетики США в штате Теннесси и Ливерморской национальной лаборатории им. Лоуренса в Калифорнии, соответственно, сохраняют первые две позиции в списке. Оба получают свою вычислительную мощность от процессоров Power 9 и графических процессоров NVIDIA V100. Система Summit немного улучшила результаты по сравнению с шестью месяцами ранее, показав рекордные 148,6 Пфлопс, в то время как система номер два Sierra остаётся неизменной на уровне 94,6 Пфлопс.

Далее идёт Sunway TaihuLight — система, разработанная Китайским национальным исследовательским центром параллельной вычислительной техники и технологий (NRCPC) и установленная в Национальном суперкомпьютерном центре в г. Уси, занимает третье место с производительностью 93 Пфлопс. Она оснащена более десяти миллионами процессорных ядер SW26010.

Под номером четыре находится суперкомпьютер Tianhe-2A (Milky Way-2A), разработанный Китайским национальным университетом оборонных техно-

логий (NUDT) и развёрнутый в Национальном суперкомпьютерном центре в Гуанчжоу. Он использовал комбинацию процессоров Intel Xeon и Matrix-2000 для достижения результата HPL 61,4 Пфлопс.

Frontera — единственный новый суперкомпьютер в первой десятке, он достиг своего пятого места с производительностью 23,5 Пфлопс. Эта система Dell C6420, работающая на процессорах Intel Xeon Platinum 8280, установлена в Техасском центре передовых вычислений Техасского университета.

Производительность компьютеров	Год	Примечание
Флопс (10^5)	1941	
Килофлопс (10^3)	1949	
Мегафлопс (10^6)	1964	
Гигафлопс (10^9)	1987	
Терафлопс (10^{12})	1997	
Петафлопс (10^{15})	2008	
Эксафлопс (10^{18})	2020?	Ожидалось в 2016 г.
Зеттафлопс (10^{21})	?	
Йоттафлопс (10^{24})	?	

Среди современных российских НРС лидирует "Ломоносов-2", занимающий 93-е место рейтинга TOP-500 с производительностью 2,478 Пфлопс. Кроме него, в TOP-500 на 365-м месте находится суперЭВМ главного вычислительного центра Росгидромета.

Создатели рейтинга отмечают, что в списке растёт доля китайских систем. Уже год назад китайские суперкомпьютеры занимали 45 % рейтинга. Системы из США занимают 22 %, но в среднем они более производительны. Если считать именно по суммарной производительности, то США лидирует с 38 %, а Китай находится на втором месте с 31 %. Также в списке лидеров по числу суперкомпьютеров Япония, Великобритания, Франция, Германия и Ирландия. У остальных стран в списке менее десяти систем.

Важно отметить, что сейчас в списке 429 суперкомпьютеров, которые демонстрируют свыше 1 Пфлопс в тесте Linpack, хотя менее года назад таких систем было лишь 272.

Если говорить о производителях систем, лидером остаётся Lenovo, которая создала 140 суперкомпьютеров. Также среди лидеров есть Sugon (57), Cray (49), HPE (46), Bull (22), Fujitsu (15), Huawei (14), Dell EMC (13) и IBM (12).

Вы спросите, а что же наши? В настоящее время Россию опережают в рейтинге не только явные технологические лидеры, но ещё десяток стран. Для сравнения, Китай представлен в рейтинге 219 суперкомпьютерами, США — 116 (уступая по числу, США обгоняют Китай по средней мощности), далее следуют Япония (29), Франция (19), Великобритания (18), Германия (14), Ирландия (13 одинаковых кластеров

Lenovo C1040). Прежде чем что-то иметь, надо ставить соответствующие задачи, в том числе и по цифровой экономике. К тому же в своё время ответственная компьютерная индустрия получила ряд чувствительных ударов, начиная с "лженауки кибернетики" в 50-е и заканчивая "мы всё купим" в 90-е годы. Кто их нанёс? Люди, не обладающие, очевидно, стратегическим и инженерным мышлением. Как результат — на рубеже тысячелетий Россия, где был всего один крупный центр МСЦ РАН, практически не была представлена в суперкомпьютерной индустрии. Однако ближе к середине 2000-х в отрасли произошёл перелом, и российские системы начали появляться в Top-500.

Сегодня по числу суперкомпьютеров в Top-500 Россия делит позиции с Финляндией, Испанией, Швецией и Тайванем. Правда, в этих странах суперкомпьютеры помощнее, но это не так важно, в этой отрасли ситуация довольно быстро меняется. Вопрос только в том, когда она изменится к лучшему в России?

Объём глобального рынка высокопроизводительных НРС-серверов продолжает расти, причём за прошлый год сумма продаж достигла 14 млрд долл. Эти серверы, разумеется, произведены не у нас, но, очевидно, этим тоже надо заниматься в процессе импортозамещения. На суперкомпьютеры в мире потрачено 5,4 млрд долл. Не будем забывать, что НРС — флагман развития новых технологий, но подчас приходится слышать, что отечественному суперкомпьютерингу не хватает задач. С другой стороны, если уж "погружаться" в НРС, то следует понимать, что по закону Мура вычислительная мощность компьютеров удваивается лишь за полтора года, а новые задачи могут появляться непрерывно, причём решение одних порождает возникновение следующих, более сложных. Растущая доступность вычислительных мощностей также стимулирует стремительное развитие технологической и конкуренции между ИТ-компаниями.

Конечно, вряд ли Россия в обозримом будущем займёт верхнюю часть рейтинга. Создание НРС на уровне приведённой выше первой пятёрки требует огромного бюджета. К примеру, стоимость суперкомпьютера Summit оценивалась примерно в 200 млн долл., исходя из чего 1 Пфлопс его мощности стоит около 1 млн долл. К тому же России необходимо строить новые суперкомпьютеры на основе отечественных технологий и процессоров, чтобы развивалась вся индустрия и связанная с НРС экосистема. Конечно, для этого требуются большие инвестиции. Но ведь это инвестиции в будущее страны.

По материалам cnews.ru,
ixbt.com, tadviser.ru, parallel.ru,
intel.ru, hpe.com, habr.com,
liberatum.ru, huawei.com, itweek.ru,
itfollow.ru, dic.academic.ru,
top500.org