

СУПЕРКОМПЬЮТЕРЫ – ГОНКА ВЫЧИСЛИТЕЛЬНЫХ ТЕХНОЛОГИЙ

Введение

Обзор посвящен актуальной на данный момент теме. Стремительное развитие технологий высокопроизводительных вычислений (НРС) напоминает известную фразу музыканта-виртуоза Ференца Листа, написанную им на полях партитуры: «быстрее, еще быстрее, быстро как только возможно... и еще быстрее». Нормативы производительности для суперкомпьютера менялись неоднократно. В 70-е суперкомпьютеры выполняли сотни миллионов операций в секунду над числами с плавающей точкой (flops, флопс). В середине 80-х был достигнут максимум в 1 гигафлопс, в конце 90-х – 1 терафлопс, а во второй половине 2000-х – 1 петафлопс. Недавно введен в эксплуатацию компьютер с пиковой производительностью в 188 петафлопс. К первой половине 2020-х несколько стран анонсировали выпуск проектов суперЭВМ нового поколения с 1 эксафлопс, а в Китае уже выпущены первые прототипы. Планка поднята очень высоко, но гонка вычислительных технологий продолжается!

Термин «суперкомпьютер»

Суперкомпьютер (Supercomputer, СуперЭВМ, Супервычислитель) — специализированная вычислительная машина, значительно превосходящая по своим техническим параметрам и скорости вычислений типовые компьютеры.

Строгого определения понятия «суперкомпьютер» не существует. Вокруг этого термина всегда велось много полемики. Шутливая классификация Гордона Белла и Дона Нельсона предлагала считать суперкомпьютером любой компьютер, весящий более тонны. Кто-то даже предлагал считать суперкомпьютером компьютер с ценой от миллиона долларов. Но, так или иначе, ни производительность, ни цена, на даже архитектура не являются однозначными критериями для отнесения вычислителя в ранг суперЭВМ.

В общеупотребительный лексикон термин «суперкомпьютер» вошёл благодаря распространённости компьютерных систем Сеймура Р. Крэй (Seymour R. Cray) таких как: CDC 6600, CDC 7600, Cray-1, Cray-2. Крэй и его команда (сначала в составе компании Control Data Corporation, а затем уже после выделения в отдельную компанию Cray Research) разрабатывали вычислительные системы, которые становились основными вычислительными средствами для крупных

правительственных, промышленных и академических научно-технических проектов.

Cray Первый

Одним из первых суперкомпьютеров считается *Cray-1*, созданный в 1976 году. С помощью поддержки векторных операций эта супер-ЭВМ с тактом 12,5 нс (80 МГц) достигала по тем временам невероятной производительности в 133 Мфлопс.

В архитектуре Cray-1 были использованы интегральные микросхемы, которые давали такую плотность упаковки логических элементов при высокой надежности, которую невозможно было достичь с помощью транзисторов. Компьютер быстро выполнял и скалярные и векторные вычисления. Обеспечение максимальной производительности достигалось за счет специальных процессорных устройств (сопроцессоры, скалярные и векторные процессоры, независимые процессоры ввода-вывода и пр.); сверхбыстрой памяти, векторных регистров команд, специализированного периферийного оборудования. Центральный процессор Cray-1 состоял из 500 печатных плат, на каждой из которых с обеих сторон располагалось по 144 микросхемы, которые охлаждались фреоном.

Уникальность инженерной конструкции вычислительного комплекса заключалась в том, что для лучшего охлаждения и циркуляции фреона в охлаждающей системе центральный процессор был выполнен в виде «башни» с 12 колоннами, составленными в форме подковы (или в виде буквы «С»), а система охлаждения была расположена в ее основании.

В 1978 году для Cray-1 был выпущен стандартный пакет программного обеспечения, состоящий из трех главных продуктов: операционная система COS (позднее UNICOS); Cray язык ассемблера (CAL); Cray Fortran (CFT), первый компилятор Fortran с автоматической векторизацией.

Историческая хроника суперЭВМ

Конечно, к первым сверхпроизводительным компьютерам можно отнести и разработанный учеными из Пенсильванского университета в середине 40-х «электронный мозг» ENIAC, мощностью 500флопс. Или вспомнить про созданную в 1952 году в нашей стране БЭСМ и ее известного потомка БЭСМ-6. И все-таки, эпоха «настоящих суперов» началась в середине 70-х ярким появлением векторно-конвейерного Cray-1. В тот момент его конкурентами

являлись такие машины, как START-100 и Cyber 76 компании CDC, ILLIAC IV компании Burroughs, IBM 70/195, STARAN компании Goodyear Aerospace и Hypercube компании IMS Associates.

В 80-е годы появилось много небольших компаний, занимающихся созданием высокопроизводительных компьютеров, однако затем большинство из них оставили эту сферу деятельности. На сегодняшний день суперкомпьютеры являются уникальными вычислительными системами, разрабатываемыми известными игроками компьютерного рынка, такими как IBM, Hewlett-Packard, Intel, NEC, Fujitsu и, конечно же, ставшим легендарным брендом «Cray».

Большинство первых суперкомпьютеров оснащались векторными процессорами, и несколько параллельно работающих векторных процессоров практически стало стандартным суперкомпьютерным решением. Конец 80-х и начало 90-х годов охарактеризовались сменой магистрального направления развития суперкомпьютеров от векторно-конвейерной обработки к большому и сверхбольшому числу параллельно соединённых скалярных процессоров. Массивно-параллельные системы стали объединять в себе сотни и тысячи отдельных процессорных элементов, причём ими могли служить не только специально разработанные, но и серийные процессоры.

В конце 90-х годов высокая суперкомпьютерных решений и нарастающая потребность в доступных вычислительных ресурсах привели к широкому распространению компьютерных кластеров. Эти системы характеризует использование отдельных узлов на основе дешёвых и широко доступных компьютерных комплектующих для серверов при помощи мощных коммуникационных систем и специализированных программно-аппаратных решений. Кластеры довольно быстро заняли достаточно большой сегмент суперкомпьютерного рынка, обеспечивая высочайшую производительность при минимальной стоимости решений.

В настоящее время основной акцент делается на увеличение числа микропроцессоров, работу с распределенной памятью, повышение степени параллелизма программ, применение графических ускорителей, создание суперкомпьютерных платформ с предоставлением облачных сервисов, а также, разумеется, поиск новых системно-технических решений.

Архитектуры супервычислителей

Далее мы будем говорить о современных *архитектурах высокопараллельных многопроцессорных вычислительных систем*. Основной принцип параллельной

обработки данных означает одновременное выполнение нескольких операций. При этом традиционно принято различать два способа обработки: параллельную и конвейерную (постадийную).

Параллельные векторные системы (PVP). Основным признаком PVP-систем является наличие специальных векторно-конвейерных процессоров, в которых предусмотрены команды однотипной обработки векторов данных, эффективно выполняющиеся на конвейерных функциональных устройствах. Как правило, несколько таких процессоров работают одновременно над общей памятью в рамках многопроцессорных конфигураций. Либо несколько таких узлов могут быть объединены с помощью коммутатора. Классическим примером компьютеров с такой архитектуры являются большинство машин Cray.

Симметричные мультипроцессорные системы (SMP) с общей памятью. Система состоит из нескольких однородных процессоров и массива общей памяти (обычно из нескольких независимых блоков). Все процессоры имеют равноправный доступ к памяти с одинаковой скоростью. Процессоры подключены к памяти либо с помощью общей шины (Front Side Bus), либо с помощью коммутатора. Аппаратно поддерживается когерентность кэшей. Для SMP-систем существуют сравнительно эффективные средства автоматического распараллеливания. Преимущества: программирование в универсальной модели общей памяти не требует какого-либо дополнительного кода или существенных изменений кода.

Массивно-параллельные системы (MPP) с распределенной памятью. Система состоит из однородных вычислительных узлов, включающих: один или несколько центральных процессоров; локальную память (прямой доступ к памяти других узлов невозможен); коммуникационный процессор или сетевой адаптер; иногда - жесткие диски и/или другие устройства ввода/вывода. Большинство массивно-параллельных систем создаются на основе мощных процессоров с архитектурой RISC (с ограниченным набором команд). Узлы связаны через коммуникационную среду. Программирование в рамках модели передачи сообщений (Message Passing Interface). Преимущества: гибкая масштабируемость, оптимальность конфигурации.

Системы с неравномерным доступом к памяти (NUMA). Система состоит из базовых модулей, состоящих из небольшого числа процессоров и блока локальной памяти. Модули объединены с помощью высокоскоростного коммутатора. Поддерживается единое адресное пространство, аппаратно поддерживается доступ к удаленной памяти, т.е. к памяти других модулей. При этом, доступ к локальной памяти в несколько раз быстрее, чем к удаленной.

Если на аппаратном уровне поддерживается когерентность кэшей во всей системе, говорят об архитектуре cc-NUMA. Примером машин с такой архитектурой является суперкомпьютер ASCI.

Кластерные системы. Кластер - это набор узлов, объединенных 2D или 3D сетью. Вычислительный узел (computational node) формируется из нескольких процессоров, традиционных или векторно-конвейерных, и общей для них памяти. Также внутренняя структура узлов может быть одинаковая (гомогенные кластеры) или разная (гетерогенные кластеры). Обычно кластер имеет как минимум один головной узел и отдельные узлы для файловой системы. Для связи узлов используется одна из стандартных сетевых технологий (Fast/Gigabit Ethernet, Myrinet) на базе шинной архитектуры или коммутатора. Плюсы кластерных систем — повышенная отказоустойчивость и более эффективная масштабируемость. По кластерному принципу построены некоторые модели суперкомпьютеров от IBM.

Правда для современных задач требуется иная архитектура, которую можно назвать *дата-центричной* (data-centric), то есть ориентированная не только на скорость счета, но и на эффективность работы с большими объемами данных.

Оригинальным направлением в архитектуре суперЭВМ являются *квантовые вычислительные системы* — устройства, использующие явления квантовой суперпозиции и квантовой запутанности для передачи и обработки данных. Такие устройства оперируют *кубитами* (квантовыми битами), которые могут одновременно принимать значение и логического ноля, и логической единицы. Поэтому с ростом количества используемых кубитов число обрабатываемых одновременно значений увеличивается в геометрической прогрессии. Фазовый кубит был впервые реализован в лаборатории Делфтского университета и с тех пор активно изучается. Квантовые компьютеры производятся, например, Microsoft и D-Wave System.

Сферы применения

В основном, суперкомпьютеры применяются в научных исследованиях, а также в областях, связанных с интенсивными вычислениями, а именно:

- физика плазмы и статическая механика;
- молекулярная и атомная физика;
- газовая динамика и теория турбулентности;
- астрофизика, космические исследования;

- метеорология, изучение и моделирование климатических, атмосферных явлений, в том числе долгосрочный прогноз погоды;
- ядерная энергетика, моделирование ядерных испытаний;
- квантовая химия, химия твердого тела;
- молекулярная динамика, теория поверхностных явлений;
- генетика, анализ и расшифровка ДНК;
- сейсморазведка, обработка сейсмических данных;
- геопространственная разведка в нефте- и газодобывающей промышленности;
- моделирование прототипов и креш-тестов в автомобилестроении;
- проектирование сложных электронных устройств;
- фармакология и синтез новых лекарственных материалов;
- обеспечение государственной безопасности, в том числе кибербезопасности, анализ военно-политических процессов;
- имитационное, ситуационное моделирование и поддержка принятия решений;
- визуализация и виртуальные среды, распознавание изображений, рендеринг, речевой синтез;
- искусственный интеллект;
- анализ больших объемов данных.

Рейтинг производительности

Рейтинг Top500 производительности суперкомпьютеров составляется дважды в год по результатам выполнения *теста Linpack* - решению методом Гаусса системы линейных уравнений с плотно заполненной матрицей коэффициентов. Производительность определяется как количество "полезных" вычислительных операций над числами с плавающей точкой в расчете на 1 секунду. Число выполненных операций с плавающей точкой оценивается по формуле $2n^3/3 + 2n^2$ (n - размер матрицы, обычно $n=1000$). Постепенно данные, собранные для рейтинга, обеспечили возможность выявления технологических и архитектурных тенденций в области высокопроизводительных вычислений. (Ниже, говоря о производительности суперкомпьютеров, мы будем иметь в виду параметр полученной на тестах максимальной производительности R_{max} , в отличие от теоретической пиковой производительности R_{peak} .) Первая восьмерка лучших выглядит в данный момент следующим образом:

1.Summit (США). В июне 2018 г. в Национальной лаборатории Ок-Ридж в штате Теннесси введен в строй новый суперкомпьютер с производительностью 122,3

петафлопс. Он разработан компанией IBM совместно с производителем графических процессоров NVIDIA. Суперкомпьютер Summit состоит из 4608 серверов, на каждом из которых установлено по два 22-ядерных процессора IBM Power9, и обладает более чем 10 петабайтами оперативной памяти. Стоимость разработки около 200 млн. долларов. Вычислительная система оптимизирована для работы с искусственным интеллектом, анализа больших массивов данных, разработки новых типов материалов.

2. *Sunway TaihuLight (Китай)*. Производительность 93 Пфлопс позволила Sunway два последних года лидировать в списке. Суперкомпьютер расположен в национальном суперкомпьютерном центре в Уси, рядом с озером Тайху. Многоядерные RISC-процессоры китайского производства серии ShenWei, количество процессоров – более 40 тысяч, общее количество ядер – более 10,5 млн. Работает под управлением операционной системы Sunway Raise OS на базе Linux. Предназначение – сложные расчеты в медицине, производстве, добывающей промышленности и анализа больших данных.

3. *Sierra (США)* был создан компанией IBM по заказу правительства США. Расположен в Ливерморской лаборатории. Производительность – 71,6 петафлопс. Состоит из 4320 серверов, каждый из которых имеет по два процессора IBM Power9 и четыре ускорителя NVIDIA Tesla V100, интерконнект – Mellanox. Общее количество ядер – более 1,57 млн. Компьютер используется при решении задач ситуационного моделирования, визуализации и искусственного интеллекта.

4. *Tianhe-2A (Китай)* был сконструирован по инициативе китайского правительства. Производительность: 61,45 петафлопс. Разработанный Оборонным научно-техническим университетом Народно-освободительной армии КНР и компанией Inspur этот «Тяньхэ» («Млечный путь») состоит из 16 тысяч узлов с общим количеством ядер в 3,12 миллиона. Применяются ускорители Matrix 2000. Оперативная память этого комплекса, занимающего 720 квадратных метров, составляет 1,4 петабайт, а объем запоминающего устройства – 12,4 петабайт.

5. *ABCI (Япония)*. Разработанный компанией Fujitsu и расположенный в Национальном институте передовой науки и технологии, производительностью 19,88 петафлопс является единственным японским суперкомпьютером в первой десятке супервычислителей. Система включает 1088 серверов Fujitsu Primergy CX2570, в каждом из которых по 2 процессора Intel Xeon Gold и 4 графических процессора NVIDIA Tesla V100. Архитектура NVIDIA Volta и ускоритель Tesla V100 требуют жидкостного охлаждения (при

этом Fujitsu применила интересный подход – охлаждение горячей водой). Суперкомпьютер используется для исследований в области искусственного интеллекта, робототехники, автопилотируемых автомобилей и медицины.

6. Piz Daint Cray XC30 (Швейцария). Самый мощный суперкомпьютер в Европе производительностью (после обновлений) 19,59 петафлопс установлен в Швейцарском национальном суперкомпьютерном центре (CSCS) в Лугано. Собран на базе 12 ядерных процессоров Xeon E5-2690 и ускорителей NVIDIA Tesla P10. Всего задействовано 362 тысяч ядер. Piz Daint, названный так в честь альпийской горы, был разработан компанией Cray и принадлежит к семейству XC30. Piz Daint применяется для различных исследовательских целей, например, компьютерного моделирования в области физики высоких энергий.

7. Titan Cray XK7 (США) с производительностью 17,59 петафлопс - компьютерный кластер под маркой компании Cray находится в Национальной лаборатории Оук-Ридж. Архитектура Titan состоит из 560 тысяч процессорных ядер и ускорителей вычислений Vidia Tesla K20X. Объем оперативной памяти – 0,71 петабайт. По большей части мощности Титана используются для программы Министерства энергетики США Innovative and Novel Computational Impact on Theory and Experiment program (INCITE), а также нескольких выбранных перспективных проектов: моделирование поведения нейтронов в самом сердце ядерного реактора, прогнозирование глобальных климатических изменений на ближайшие несколько лет и др.

8. Sequoia (США). Мощнейший из суперкомпьютеров семейства IBM Blue Gene/Q (производительность: 17,17 петафлопс; 96 стоек; 1,57 миллионов процессорных ядер; оперативная память – 1,57 петабайт), расположен в США в Ливерморской национальной лаборатории. IBM разработали Sequoia для Национальной администрации ядерной безопасности (NNSA), которой требовался высокопроизводительный компьютер для вполне конкретной цели – моделирования ядерных взрывов. Мощности суперкомпьютера использовались для решения и других задач. Кластеру удалось поставить рекорды производительности в космологическом моделировании, а также при создании электрофизиологической модели человеческого сердца.

Суперкомпьютеры в нашей стране

Ломоносов-2, входящий в первую сотню рейтинга с производительностью 2,48 петафлопс, построен компанией «Т-Платформы» для МГУ им. М.В.Ломоносова и установлен в НИВЦ МГУ. Система состоит из 1280 узлов на базе процессоров Intel Skylake и ускорителей Nvidia Pascal P100. Общее количество

ядер более 64 тысяч. Управляющие, вычислительные и коммутационные модули A-Class охлаждаются горячей водой с температурой до +45°C. На суперкомпьютерах Московского государственного университета работают около 2500 пользователей из 20 подразделений МГУ, более 100 институтов РАН, более 100 университетов России, решая задачи по магнитной гидродинамике, гидро- и аэродинамике, квантовой химии, сейсмике, моделировании лекарств, геологии и науке о материалах, криптографии.

Группа компаний «РСК» установила *MBC-10П МП* на основе массивно-параллельной архитектуры RSC PetaStrem в Межведомственном суперкомпьютерном центре РАН. Система расширила вычислительные ресурсы МЦС, основу которых составляет система MBC-10П на базе кластеров «РСК Торнадо» увеличила суммарную производительность до 0,6 петафлопс.

Также достойны упоминания и совместные государственные программы союзного государства России и Беларуси по развитию суперкомпьютеров «СКИФ» (2000—2003гг, 2007—2010гг, 2011—2014гг). По состоянию на 2010 год, 8 компьютеров России, попадавших в Top500, относились к семейству СКИФ. Несколько систем входит в Top50 суперкомпьютеров СНГ, включая СКИФ МГУ и СКИФ Аврора, разработанный при участии Института Программных Систем РАН.

Статистика распределения в мире

- Китай – 206;
- США – 124;
- Япония – 36;
- Великобритания – 22;
- Германия – 21;
- Франция – 18;
- Нидерланды – 9;
- Ю.Корея – 7;
- Ирландия – 7;
- Канада – 6;
- Австралия – 5;
- Индия – 5;
- Италия – 5;
- Россия – 4;
- другие страны – 25.

(Представлены данные рейтинга Top500 на июнь 2018г.)

Заключение

Тенденции дня сегодняшнего свидетельствуют о том, что в ближайшем будущем нас ожидает этап концентрации информационных ресурсов в больших вычислительных центрах или центрах обработки данных нового поколения. (Подробнее – см. статью в журнале ЦЭ [«Центры обработки данных нового поколения – взгляд на горизонт»](#)).

Поэтому, не менее актуальным является проблема быстрой интеллектуальной обработки больших массивов данных, с которой помогут справиться именно суперкомпьютеры. К сожалению, в этом процессе Россия пока сильно отстает от лидеров (см. выше распределение суперкомпьютеров по странам). В самое ближайшее время нам предстоит решить несколько важных задач, а именно:

- 1) Разработка отечественного компьютера с производительностью в сотни и тысячи петафлопс.
- 2) Разработка и применение новой технологии компьютерного конструирования на универсальных вычислительных элементах (совмещение функций процессора и хранения информации или «активная память» по удачному выражению основателя ВЦ РАН академика А.А. Дородницына).
- 3) Разработка отечественных высокоскоростных магистральных сетей передачи данных для объединения суперЭВМ.
- 4) Разработка и развитие средств параллельного программирования: коммуникационные интерфейсы, параллельные языки и расширения языков.