

Современные суперкомпьютеры, как системы массового параллелизма обработки данных

В.К. Левин

В стоимостном объеме мирового выпуска компьютерной продукции суперкомпьютеры составляют небольшую долю, однако уровень их производительности, определяющий возможность решения наиболее сложных вычислительных задач и обработки данных, всегда признавался, как один из стратегических показателей научно - технического потенциала. В суперкомпьютерах воплощаются основные вновь возникающие структурно - технические решения, которые затем распространяются в разнообразных компьютерах массового применения. Термин "суперкомпьютер", как определение класса техники, получил распространение в связи с выпуском вычислительных систем Gray-1, Gray-2 и им подобных с наращиванием производительности в диапазоне 0,1 - 10 млрд. оп./сек. (конец 70-х - начало 90-х гг.), хотя он применялся и ранее, а также продолжает использоваться в настоящее время. В преддверии появления к 2000 г. суперкомпьютеров производительностью триллион и более оп./с. появились термины "ультракомпьютер" и "Теракомпьютер", хотя по-прежнему используется традиционный термин "высокопроизводительная вычислительная техника" (HPC - High Performance Computers).

Повышение производительности, определяемой количеством выполняемых операций (отнесенным к интервалу времени) должно быть сбалансировано ростом объема и быстродействия памяти, интенсивности ввода - вывода данных, а также развитием программного обеспечения. Рост физического быстродействия компонентов происходит, грубо говоря, вдвое медленнее роста производительности, но зато возрастает количество компонентов в суперкомпьютере, что соответствует развитию параллелизма на всех структурных уровнях системы [1,2].

Современная физико - технологическая база компьютеров и, вообще, средств автоматизации, характеризуется, прежде всего, прогрессирующим выпуском микропроцессоров, в структурно - техническом построении которых воплощены основные достижения вышеупомянутых суперкомпьютеров 80-хх гг. Серийные микропроцессоры уже превзошли уровень 100 млн. оп./с., и выходят на уровень 1 млрд. оп./с. при тактовой частоте до 500 МГц и интеграции порядка 3-5 млн. транзисторов на кремниевом кристалле однокорпусного микропроцессора; повидимому, эти показатели в последующие годы будут повышены еще в 3-5 раз. Развитость функций современных микропроцессоров делает их достаточно пригодными как для массовых компьютеров, так и для суперкомпьютеров, в то время, как еще недавно для суперкомпьютеров создавались специфические элементы с возможно более высоким быстродействием, причем относительно небольшой объем их выпуска не стимулировал повышение уровня интеграции и снижение стоимости. Конечно, работы над элементами с высоким быстродействием продолжаются и вполне возможно освоение каких-то новых физических решений, которые могут стать достоянием перспективных суперкомпьютеров (элементы на арсениде галлия, оптические, криогенные и др.).

Структурный параллелизм разных уровней аппаратуры (микрооперации и команды процессора, объединение процессоров общей памятью и каналами обмена) определенным образом отражаются в программировании как на системном, так и на прикладном уровне. Изначально компьютеры и методы их применений (математические модели, формализация задач) развивались, как взаимоувязанные процессы. На всех этапах развития компьютеров формировался значительный задел программ и методов, которые являлись консервативным фактором по отношению к вновь возникающим техническим структурам. Приходится идти на определенные издержки, вводя стандартизацию языков программирования, человеко - машинных и аппаратно - программных интерфейсов, которые как-то обеспечивают преемственность создаваемого задела, поскольку меняются реже (как бы, более крупными ступенями), чем быстро развивающиеся конкретные аппаратные и программные модули. Вопросы формирования и внедрения концепций открытых систем выходят за рамки данной статьи, вместе с тем, они рассматриваются, как весьма важные в развитии структур параллельных вычислений и обработки данных.

Параллелизм микроопераций здесь не рассматривается, поскольку он освоен в предыдущих реализациях, хотя даже и сейчас некоторые вопро-

сы возникают вплоть до уровня прикладных программ, например, вопрос об эффективности выполнения поразрядных (побайтных или побитных) операций в современных 64 - битных микропроцессорах по сравнению с какими-либо другими компьютерами (специализированными или "старыми" 8 - битными микропроцессорами). Характерным для современных микропроцессоров является векторно - конвейерный параллелизм, т.е. совмещение выполнения нескольких подряд идущих команд программы. Обычно микропроцессоры характеризуются пиковым быстродействием (оп./с.), в 2 раза (иногда - в 3-4 раза) превышающим тактовую частоту. Поскольку многие операции занимают несколько тактов (порядка 10), имеет место совмещение 20 (и более) операций (команд). Это совмещение осуществляется настолько, насколько операции независимы, чтобы не нарушались причинно - следственные связи между командами (обработываемыми данными).

Другой проблемой эффективного использования микропроцессора является его взаимодействие с присоединяемой оперативной памятью, объем которой в процессорных узлах современных суперкомпьютеров варьируется от 8-16 Мбайт до 1-2 Гбайт. Требование большого объема оперативной памяти определяется усложнением решаемых задач и ограниченностью скорости ввода - вывода данных в процессорный узел (со стороны других процессоров и внешней памяти). Быстродействие оперативной памяти (даже если она разбита на параллельно работающие модули и имеет двухуровневую иерархию по быстродействию) не соответствует непосредственно быстродействию микропроцессора, поэтому внутри последнего ставится 1-2 уровня кэш-памяти объемом до нескольких десятков Кбайт. Для эффективной работы процессорного узла необходима достаточная степень локализации данных в кэш-памяти, чтобы ее обмен данными с присоединяемой оперативной памятью проходил в том относительно "невысоком" темпе, который соответствует возможностям памяти, причем этот обмен предпочтительно вести массивами из нескольких чисел (порядка 100 байт), располагаемых в подряд идущих адресах, т.к. именно при этом существующая организация памяти дает наивысший темп обмена.

Объединение нескольких микропроцессоров с целью суммирования производительности осуществляется двумя путями. Первый основан на взаимодействии нескольких (обычно до 8) микропроцессоров с общей оперативной памятью и считается наиболее быстродействующим, хотя и возникают издержки в связи с определенной дисциплиной разграни-

чения (перераспределения) памяти между процессорами, поддержанием так называемой когерентности кэш-памятей и др. Технически затруднительно объединять в системах с общей памятью (они часто называются симметричными мультипроцессорами - SMP) большое число процессоров.

Второй путь основан на объединении процессорных узлов через встраиваемые в них внешние каналы (линки - в транспьютерах, связи через адаптеры канал - канал в многомашинных комплексах ЕС ЭВМ и т.п.). Имеется, таким образом, некоторая межпроцессорная коммутационная среда, которая отчасти распределена по процессорным узлам, отчасти сосредоточена в специальных коммутаторных узлах системы. Распределенная коммутационная среда применена в отечественной мультипроцессорной системе МВС-100, в которой каждый процессорный узел содержит основной микропроцессор 1860 и в качестве связного процессора - транспьютер Т805 [2]. Межпроцессорный обмен организуется по некоторой дисциплине обменов сообщениями ("точка - точка") в виде массивов данных по нескольку сотен и тысяч байт; короткие сообщения пересылать невыгодно из-за затрат на настроечные процедуры (вхождение в связь, маршрутизация и проч.). Такие системы (при надлежащем построении коммутационной среды) допускают широкое масштабирование (расширяемость до сотен и тысяч процессоров) и называются системами массового параллелизма (МРР - Massively Parallel Processing); и ["несимметрия" по сравнению с системами SMP выражается в том, что взаимодействие конкретного микропроцессора со "своей" оперативной памятью осуществляется непосредственно по общепринятой схеме, а доступ к памяти других процессорных узлов, хотя и может быть непосредственно (виртуально) адресован, но реализуется по межпроцессорному тракту передачи сообщений. 10-15 лет назад усматривалась некоторая альтернативность между системами SMP и МРР в том смысле, что МРР-системы рассматривались, как однородное множество "простых" и, поэтому, дешевых процессоров, что могло, при определенных условиях, быть выгоднее, чем SMP-система из небольшого числа мощных векторно-конвейерных процессоров. В условиях современной микропроцессорной базы такого рода альтернативность себя исчерпала, однако считается, что SMP-системы предпочтительнее при небольшом числе процессоров (SMP-кластеры). Кроме того может признаваться нецелесообразным построение МРР-системы, как однородной в составе сотен и тысяч процессоров, и строится МРР-система, как объединение SMP-

кластеров.

Общесистемная организация мультипроцессорной системы обычно предполагает наличие довольно мощного хост - компьютера, через который (или несколько таких) осуществляются функции внешнего взаимодействия с мультипроцессорным (МР) массивом. Внешняя память присоединяется через хост-компьютер, а также непосредственно к МР-массиву. Естественно, что возникает функциональная (а иногда и техническая) неоднородность в виде выделения "корневых" процессоров МР-массива, через которые идет внешнее взаимодействие.

Необходима сбалансированность характеристик быстродействия процессора и объема памяти с характеристиками межпроцессорного обмена, собственно вычислительные операции должны быть, по возможности, совмещены (по времени) с обходами, для чего, в частности, обработка и обмены данными должны быть равномерно распределены по МР-массиву. Эта сбалансированность связана с характеристиками (потребностями) решаемых на суперкомпьютере прикладных задач и с используемой "технологией" системного программирования (включая разного рода виртуальные представления и режимы параллельной обработки), которые сами по себе весьма разноречивы. Привычка к сколько -нибудь определенному классу задач зачастую расценивается как специализированное построение. Есть также технико - экономический подход к определению параметров: затраты на построение собственно обработки (основной микропроцессор), оперативную память и цепи межпроцессорного обмена должны быть равны ("равнопрочная" система); затраты на МР-массив и хосткомпьютер (а также внешнюю память) д.б. также равны между собой; можно также приравнивать затраты между аппаратными и программными средствами системы. Представляется, что при таком подходе минимизация затрат будет сочетаться с универсализацией применений системы. Имеется, однако, вопрос объективной оценки вышеупомянутых затрат.

Мировая практика демонстрирует большой разброс суперкомпьютерных решений MPP и SMP - систем, что свидетельствует о продолжающемся поиске в этой области. Применительно к отечественным разработкам этого направления заслуживают внимания проблемно - ориентированные системы, как более целенаправленные и экономичные. Необходимо также вести работы по "нетрадиционным" системам с учетом меры их соответствия современной элементной базе; в этом смысле хорошие предпосылки имеют нейрокомпьютерные структуры, как своего рода си-

стемы массового параллелизма обработки данных.

В заключение следует отметить, что в современных условиях изменяется режим работы территориально распределенных вычислительных сетей. Терминалы пользователей в сетях зачастую представляют собой компьютеры с производительностью 0,1-1 млрд. оп./с., поэтому потребность в обращении к удаленному (центральному) компьютеру возникает тогда, когда требуется заметно большая производительность. Поскольку центральным компьютером может быть суперкомпьютер с массовым параллелизмом обработки, целесообразно обслуживать удаленных пользователей путем выделения некоторого числа процессоров из полного МР-массива (а не путем разделения времени в мультипрограммном режиме, как это делалось традиционно в мощных однопроцессорных или SMP-центрах). Естественно должны быть приняты меры для разграничения пользователей, устранения влияния отказов отдельных процессорных узлов и проч., чтобы обеспечить устойчивость и безопасность работы вычислительной сети.

Список литературы

- [1] В.К. Левин. Радиоэлектроника и вычислительная техника. "Радиоэлектроника". М. 1995, N 4-5, 137-141.
- [2] В.К. Левин. Высокопроизводительные мультипроцессорные системы. "Информационные технологии и вычислительные системы". М. 1995, N 1, 12-21.