

УДК 004.27

Генезис и альтернативы развития высокопроизводительной вычислительной техники

В.К.Левин

В работе охарактеризованы концепции построения и факторы развития высокопроизводительных вычислительных систем, приведены прогнозные данные.

производительность; суперкомпьютер; массово-параллельные кластеры; гибридные вычислительные модули; многопоточная обработка данных.

Представление, передача и хранение информации дискретными электрическими сигналами, электронная схемотехника, реализующая логические операции «И», «ИЛИ», «НЕ», а также триггерные счетчики импульсов – все это было известно до 1940 г. В конце 1945 г. в США была введена в действие первая электронно-ламповая цифровая вычислительная машина с программным управлением – ENIAC в составе 18 тыс. электронных ламп. Эта машина имела производительность (или, как тогда говорилось, быстродействие) порядка 100 арифметических операций в секунду. При создании последующих машин к 1950 г. концепция программного управления – так называемая «фон-неймановская» – сформировалась в довольно целостном виде. Эта концепция получила дальнейшее развитие и в настоящее время сохраняет свое значение как фундаментальная.

В развитии электронной вычислительной техники могут быть названы «поколения», «узловые» (или «переломные») моменты и драматические события. Однако, в целом, прошедший 65-летний период создания и применения этой техники выглядит как бурно протекающая эволюция, определяемая высоким темпом роста производительности и расширением применений компьютеров. Производительность этой техники растет за счет повышения физического быстродействия и количества компонентов, что связано с удешевлением, миниатюризацией, сокращением энергопотребления компонентов и параллелизмом их работы в структуре компьютера.

Как и во многих других областях, развитие в сфере вычислительной техники и ее применений можно представлять в виде восходящей спирали, поскольку на каждом новом этапе, наряду с ростом количественных показателей, происходит качественное обновление, в котором можно усматривать аналогии с тем, что было на предыдущих этапах.

В качестве компьютерной базы электронные лампы на рубеже 1960 г. сменились на транзисторную схемотехнику (компьютеры 2-го поколения), а на сегодня – прежде всего, многоядерные микропроцессоры и СБИС памяти. В последнее 10-летие в построении наиболее мощных вычислительных средств – суперкомпьютеров доминируют массово-параллельные кластеры (MPP-структуры), строящиеся как объединение серийных вычислительных модулей и сетевых средств. Этому предшествовали системные структуры SIMD и SMP, которые сейчас, по существу, вошли в структуру многоядерных микропроцессоров.

Производительность наиболее мощной вычислительной установки возростала, в среднем, 10-кратно за пятилетие, а в течение последних 25 лет – более чем 20-кратно, т.е. 1000-кратно за 11 лет (рис. 1) [1]. Появившаяся в 2008 г. вычислительная система-кластер Roadrunner (ф. IBM, США) имеет производительность 1 Petaflops (10^{15} – квадриллион оп./с) по тесту Linpack. В 2010 г. зарегистрировано 7, а в 2011 г. – 10 суперкомпьютеров, имеющих производительность более 1 Pflops. В России также активизируются работы этого направления [2, 5–8]. В прогнозных оценках предсказывается выход на уровень производительности 100 Pflops в 2015–16 гг. и 1 Exaflops (10^{18} оп./с) – в 2018–20 гг. (подразумевая оценки производительности сопоставимыми способами, поскольку тест Linpack может быть чем-то заменен).

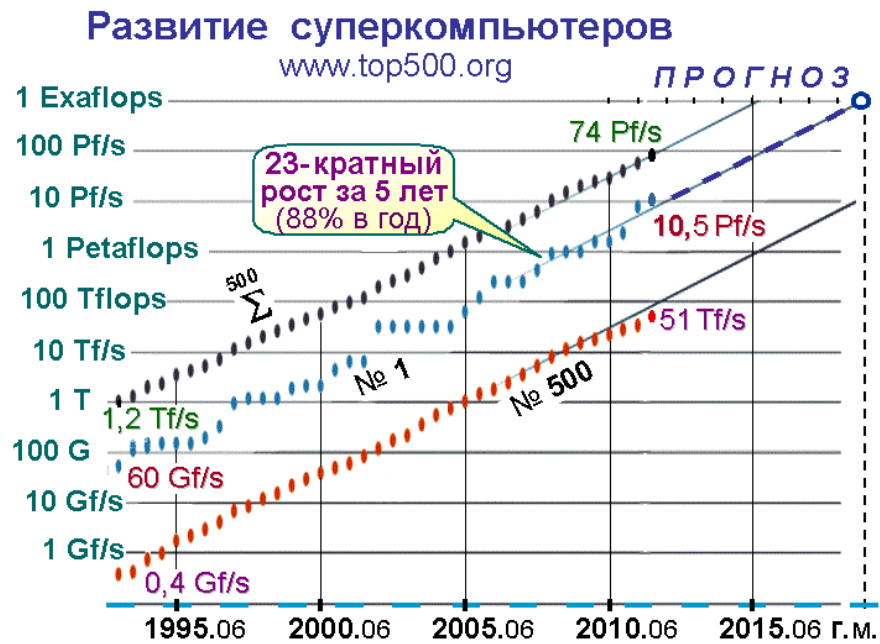


Рис. 1. Развитие суперкомпьютеров

В «стержневом русле» развития вычислительной техники, имеются многочисленные ответвления, некоторые из них выделились как самостоятельные направления. Естественно, это развитие сопровождалось рассмотрением разного рода альтернативных подходов и решений [3–8].

Требования повышения производительности и доступности суперкомпьютеров определяются возрастанием объемов вычислений и усложнением методов применений по мере развития компьютерного моделирования сложных процессов и объектов. Происходит «интеллектуализация» машинной обработки данных, сопоставление функций автоматов и биологических организмов.

В качестве основополагающих альтернатив можно назвать:

- сопоставление (или сочетание) универсальности и специализации систем – мягкая и жесткая (узкая) специализация, «нетрадиционные» структуры и устройства;
- усложнения структур – крупноблочные либо «мелкозернистые» структуры; однородные и гетерогенные; конвейерные и непосредственно-параллельные.

Функционально-специализированные структуры могут применяться для непосредственного моделирования процессов в виде заказной или перенастраиваемой (реконфигурируемой, структурно-программируемой) аппаратуры под конкретные задачи или классы задач.

В настоящее время все более широко используются гибридные вычислительные модули (процессор + сопроцессор) в рамках MPP-системы. Совершенствуются скалярная,

векторная, многопоточковая (multithreading) обработки, графические сопроцессоры (ГП) реконфигурируемые структуры (на базе ПЛИС – FPGA) и др. Можно вспомнить, что специализированные сопроцессоры, а на их основе гетерогенные и специализированные системы были в вычислительной технике 2-го поколения; позже появились *array-processors*.

В качестве сопутствующих и, отчасти (или на первых порах), специализированных направлений в вычислительной технике (информатике) можно назвать:

- аналоговые и аналого-цифровые системы;
- систолические, потоковые, ассоциативные структуры;
- средства машинного «интеллекта»;
- нейрокompьютеры;
- устройства нечеткой логики; вероятностные автоматы;
- работоспособные построения при «ненадежных» базовых элементах;
- использование «нетрадиционной» физико-технологической базы – оптические, магнитные, одноэлектронные, квантовые, химические, биологические элементы, среды и процессы;
- развивающиеся нанотехнологии.

В современных суперкомпьютерах продолжает развиваться многоуровневый параллелизм в иерархических (вертикальных) структурах с горизонтальными связями. При этом форсируются межпроцессорные обмены за счет повышения скорости (в потоке данных) и снижения латентности (задержки при доступе к данным). Используются сочетания структур коммутаций типа дерево, кольца, гиперкуб и др. Организуется виртуальная общая оперативная память (общая внешняя память возникла в мощных машинах 2-го поколения). Имеет место разнообразие подходов в использовании ГП и ПЛИС в сочетании с микропроцессорами общего применения.

Библиотеки и пакеты программ стали преемственным товарным продуктом. В целом, программирование усложняется (многие «старые» вопросы программирования обостряются в новое время). Программное обеспечение в условиях распараллеливания вычислений эволюционирует от описания процессов параллельной обработки к попыткам автоматизации распараллеливания программ.

После появления термина «суперкомпьютер»^{*)} и осознания трудностей применения таких систем появились «мини-суперкомпьютеры» – в том числе, для начального освоения параллельной обработки. Сейчас в этой роли выступают серверы серийного производства с относительно небольшим числом процессоров.

При создании вычислительной техники, можно сказать, изначально применялась автоматизация проектирования. Сейчас – это развивающиеся системы (САПР). Они используются также для настройки ПЛИС – это синтез: схемотехника + САПР + программирование.

Вся история вычислительной техники связана с обеспечением надежности в условиях отказов при наращивании объема аппаратуры.

На всех этапах обновляются подходы к размещению и конструктивному оформлению аппаратуры, энергообеспечению и теплоотводу (новое сопоставимо с решениями многолетней давности).

По мере развития вычислительной техники возникали новые возможности для построения встраиваемых систем и систем реального времени.

Видоизменяется и совершенствуется организация процессов обработки: концепция «клиент-сервер», роль main frame, удаленный доступ к вычислительным ресурсам и др. В 1959 г. А.И.Китов (1920–2005), обращаясь в правительственные инстанции, выдвинул предложения о построении сети вычислительных центров (ВЦ) – объединении ресурсов для решения больших задач, «не помещающихся» в одной вычислительной машине (вплоть до создания единой государственной сети – последующие проекты ОГАС, ЕГСВЦ). В том же направлении работал в Сибирском отделении РАН А.П.Ершов (1931–1988); ссылаясь на работы, проводившиеся в США, он пропагандировал концепцию удаленного доступа абонентов к сети ВЦ и с 1966 г. развернул работы по абонентским терминалам и операционным системам для режима разделения времени в ВЦ коллективного пользования (использовалось название АИСТ –автоматизированные информационные станции). Работы этого направления велись также под руководством В.М.Глушкова Специализированные вычислительные комплексы, обрабатывающие данные от удаленных источников и передающие результаты потребителям (объектам управления) строились еще на базе ламповой техники.

^{*)} В 1929 г. в газете New York World термин Super Computing применен в связи с мощным электромеханическим табулятором ф. IBM (<http://ru.wikipedia.org>). В публикациях по

На сегодня для крупных ВЦ режим обслуживания удаленных пользователей стал повсеместным. Сети компьютеров, которые замыслились на базе вычислительной техники 1–2 поколений, представлены Grid-технологиями и, так называемыми, облачными вычислениями с использованием мощных центров обработки и хранения данных (ЦОД). Вместе с тем, для многих применений вычислительная техника может оказаться относительно недорогой (по сравнению с объектами которые с ее помощью обслуживаются). Поэтому зачастую целесообразно использовать сосредоточенные вычислительные установки требуемой производительности непосредственно там, где надо, минимизируя затраты и организационные трудности применения средств высокоскоростной связи и защиты передаваемой информации. В общем, в современных системах имеет место сочетание сетевых концепций с централизацией обработки и хранения данных.

Необходимо сознавать и оценивать ограничивающие факторы в построении суперкомпьютеров, формировать технико-экономические показатели и критерии оценок системных построений. В той или иной форме соблюдается сбалансированность разного рода затрат в сферах собственно вычислительной техники, в её применениях и в физико-технологической и производственной базе, – эти затраты подразделяются на этапы разработки (НИР, ОКР), изготовления и эксплуатации систем. В пределах собственно системы должно быть обосновано соотношение различных аппаратных и программных средств (обработка, хранение и передача данных, управление, программное обеспечение системного и прикладного уровней); это относится также к системно-сетевому уровню (централизация, либо распределённость ресурсов). Проявляются факторы развития и консерватизма в таких вопросах как масштабирование (расширяемость) систем, стандарты, преимственность задела. Приходится также учитывать относительно быстрое моральное старение техники (создание большого суперкомпьютера может занимать немалое время, и для решения вновь возникающих задач он может оказаться устаревшим вскоре после его создания, т.к. обновляются поколения микроэлектроники и системно-структурные концепции). Современные мощные вычислительные системы строятся на передовой компонентной базе массового выпуска (используются те же микропроцессоры, что и в массовых компьютерах; создание каких-то уникальных и малотиражных компонентов (на сегодня – микроэлектронных) признается неоправданным. Однако в будущем могут проявляться и другие подходы (использоваться некоторые уникальные компоненты и проч.).

электронной вычислительной технике термин *суперкомпьютер* стал широко использоваться после 1965 г.

На сегодня упомянутая выше альтернатива – крупноблочные либо «мелкозернистые» структуры – применительно к микропроцессорам проявляется в том, что в списке Top 500 сосуществуют системы на базе микропроцессоров с тактовой частотой более 3 ГГц и с тактовой частотой менее 2 ГГц. Последние позволяют строить системы с меньшим энергопотреблением, но с увеличенным количеством ядер, а это усложняет программирование. В еще большей мере усложняется программирование для гибридных систем с ГП и процессорами типа Cell (в списке Top 500 от ноября 2011 г. они занимают 39 мест). Все это свидетельствует об острой конкуренции в области вычислительной техники и микроэлектроники.

Развертывание систем (наращивание мощности и освоение) осуществляется поэтапно. Системы формируются как гетерогенные с функционально-ориентированными подсистемами и работают в сетевой среде с рабочими станциями пользователей.

Формирование вышеупомянутой фон-неймановской концепции как основы электронной цифровой вычислительной техники несомненно опиралось на имевшийся к тому времени опыт использования численных методов решения вычислительных задач с применением электромеханической вычислительной техники. В дальнейшем, с ориентацией на электронику, численные методы интенсивно и всесторонне развивались. Вместе с тем сознается необходимость адаптации и взаимосвязи, с одной стороны, математических методов и программ, а с другой – специфики построения технических структур (параллельная обработка и др.) с соответствующей организацией прохождения потока задач (очередность, приоритеты).

В начавшемся десятилетии предстоит выяснить, как будут создаваться и применяться Exaflops-ные системы – эволюционно или революционно [7–9]. Будут развиваться и осваиваться проблемно-ориентированные территориально-распределенные гетерогенные системно-сетевые построения на прогрессирующей элементной базе.

Литература

1. <http://www.top500.org> (списки и обзор 500 наиболее мощных вычислительных установок мира, начиная с 1993 г.).
2. <http://supercomputers.ru/> (списки и обзор 50 наиболее мощных вычислительных установок России и стран СНГ, начиная с 2004 г.).
3. Левин В.К. Очерк становления Единой системы ЭВМ.
http://www.computer-museum.ru/histussr/es_levin.htm

4. Левин В.К. Структурно-технические характеристики и направления развития высокопроизводительных вычислительных систем. / Электронная вычислительная техника // Под ред. В.В.Пржиялковского. – М.: «Радио и связь», – 1988, – вып. 2, – С. 4–16.
5. Левин В.К. Радиоэлектроника и вычислительная техника.
// Радиотехника, № 4–5, 1995, С.137–141.
6. Левин В.К. Этапы развития суперкомпьютеров.
http://www.computer-museum.ru/histussr/super_phase_0.htm
7. Забродин А.В., Левин В.К. Суперкомпьютеры – состояние и развитие. // Автоматика и телемеханика, 2007, №5, С. 4–7.
8. Лацис А.О. и др. Гибридный суперкомпьютер К-100: что дальше?
http://www.hpc-platform.ru/tiki-download_file.php?fileId=77
<http://www.kiam.ru/MVS/resourses/k100.html>
9. Многоточие Стерлинга (беседа В.Воеводина и И.Левшина с Т.Стерлингом)
// Суперкомпьютеры, 2010, №3, С. 17–20.
http://www.supercomputers.ru/index.php?option=com_k2&view=item&id=94:многоточие-стерлинга&Itemid=2.

Левин Владимир Константинович; профессор Федерального государственного унитарного предприятия «Научно-исследовательский институт «КВАНТ»»; д.т.н., академик РАН, член-корр. РАН; тел.: 8(499)745-73-20, e-mail: levin@rdi-kvant.ru,

