

Нижегородский государственный университет им. Н.И. Лобачевского  
Факультет вычислительной математики и кибернетики

**Образовательный комплекс  
«Введение в принципы функционирования и  
применения современных мультиядерных  
архитектур (на примере Intel Xeon Phi)»**

**Лекция №1  
Обзор архитектуры современных многоядер-  
ных процессоров**

---

*Сиднев А.А., Сысоев А.В.*

*При поддержке компании Intel*

Нижний Новгород  
2013

## Содержание

<b>ВВЕДЕНИЕ .....</b>	<b>3</b>
<b>1. ПАРАЛЛЕЛИЗМ КАК ОСНОВА ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛЕНИЙ .....</b>	<b>6</b>
1.1. СИММЕТРИЧНАЯ МУЛЬТИПРОЦЕССОРНОСТЬ .....	7
1.2. ОДНОВРЕМЕННАЯ МНОГОПОТОКОВОСТЬ.....	11
1.3. МНОГОЯДЕРНОСТЬ .....	13
<b>2. МНОГОЯДЕРНОСТЬ – ДВА, ЧЕТЫРЕ, ВОСЕМЬ – КТО БОЛЬШЕ? .....</b>	<b>16</b>
2.1. ПРОЦЕССОРЫ INTEL® CORE™ И INTEL® XEON® .....	16
2.2. ПРОЦЕССОРЫ AMD PHENOM™ И AMD OPTERON™ .....	19
2.3. ПРОЦЕССОРЫ IBM POWER7.....	20
2.4. ПРОЦЕССОРЫ POWERXCELL™ 8i.....	21
2.5. ПРОЦЕССОРЫ SUN ULTRASPARC T1, SUN ULTRASPARC T2, SPARC T3, SPARC T4.....	23
<b>3. УСКОРИТЕЛИ ВЫЧИСЛЕНИЙ.....</b>	<b>26</b>
3.1. УСКОРИТЕЛЬ CLEAR SPEED™ ADVANCE™ E710 .....	26
3.2. GPU-УСКОРИТЕЛИ NVIDIA® TESLA K20X.....	28
3.3. СОПРОЦЕССОР INTEL® XEON PHI™ 5110P.....	29

## Введение

«Citius, Altius, Fortius»<sup>1</sup> – девиз Олимпийских игр современности, как ни к какой другой области, применим к вычислительной технике. Воплощение в жизнь не раз видоизменявшего свою исходную формулировку, но до сих пор действующего эмпирического закона, сформулированного в 1965 г. Гордоном Муром, похоже, стало «делом чести» производителей аппаратного обеспечения. Из всех известных формулировок этого закона точку зрения потребителя/пользователя наилучшим образом отражает вариант: «производительность вычислительных систем удваивается каждые 18 месяцев». Мы сознательно не использовали термин «процессор», поскольку конечного пользователя вовсе не интересует, кто обеспечивает ему повышение мощности: процессор, ускоритель, видеокарта – ему важен лишь сам факт роста возможностей «за те же деньги».

Закон Мура в общем случае гласит, что за небольшой период времени происходит удвоение количества транзисторов на кристалле<sup>2</sup>. Величина этого периода обычно составляет 1-2 года и определяется уровнем развития промышленности в конкретный момент времени. Закон Мура достигается за счёт более плотного размещения транзисторов, т.к. линейное расстояние между электродами транзисторов уменьшается (происходит уменьшение техпроцесса, определяющего размер транзисторов). В результате тепловыделение и энергопотребление транзисторов уменьшается. Как следствие, становится возможным размещение большого количества транзисторов на кристалле с сохранением прежних (или даже меньших) показателей тепловыделения и энергопотребления.

До настоящего момента увеличение количества транзисторов выражалось в росте тактовой частоты процессоров, увеличении размеров локальной памяти процессоров (кэш-памяти<sup>3</sup>) и усложнению логики процессоров (интеллектуальные предсказатели ветвлений, многоэтапные конвейеры, блоки векторных вычислений, возможность выполнять несколько инструкций за такт и др.). Основным способом повышения производительности было увеличение тактовой частоты процессоров. Со временем это стало невы-

---

<sup>1</sup> «Быстрее, Выше, Сильнее» – лат.

<sup>2</sup> Имеется несколько законов Мура, вот один из них: "число транзисторов в производимых чипах удваивается каждые два года". Формулировка: "производительность микропроцессоров удваивается каждые 18 месяцев", - была выдвинута Дэвидом Хаусом на основе предыдущего закона, но принято и эту формулировку считать "законом Мура".

<sup>3</sup> Первые процессоры либо не содержали кэш-память, либо их кэш-память была размером несколько КБ. Современные процессоры общего назначения имеют многоуровневую иерархию кэш-памяти, суммарным размером несколько МБ.

годно в силу физических ограничений, т.к. увеличение тактовой частоты приводит к нелинейному росту тепловыделения и потребляемой мощности<sup>4</sup>. Производители, выбрав в качестве магистрального пути развития увеличение числа ядер на кристалле, были вынуждены призвать на помощь разработчиков программного обеспечения. Старые последовательные программы, способные использовать лишь одно ядро, теперь уже не будут работать быстрее на новом поколении процессоров «задаром» – требуется практически повсеместное внедрение программирования параллельного.

Кроме представленной выше известна и другая формулировка закона Мура: «доступная (человечеству) вычислительная мощность удваивается каждые 18 месяцев». Зримое свидетельство этого варианта формулировки – список Top500 [106] самых высокопроизводительных вычислительных систем мира, обновляемый дважды в год. В ноябре 1993 г. был представлен 2-ой список Top500, содержащий одну вычислительную систему с производительностью больше 100 гигафлопс (Numerical Wind Tunnel производства Fujitsu). Барьер в один терафлопс был преодолен системой ASCI Red [111] производства компании Intel в 1997 г. В июне 2005 система IBM BlueGene/L имела производительность более 100 терафлопс, а уже в 31-м списке Top500 (июнь 2008 г.) впервые в истории был преодолен петафлопный порог производительности – суперкомпьютер Roadrunner [112] производства компании IBM показал на тесте LINPACK 1,026 петафлопс (всего за 15 лет пик мощности вырос на четыре порядка). А суммарная мощность систем, представленных в 31-м списке Top500, составила 11,7 петафлопс. Много это или мало? Если взять за основу, что реальная производительность хорошей «персоналки» на четырехъядерном процессоре составляет порядка 20 гигафлопс, то весь список Top500 будет эквивалентен половине миллиона таких персоналок. Очевидно, что это лишь вершина айсберга. По данным аналитической компании Gartner, общее число используемых в мире компьютеров превысило в 2008 г. 1 миллиард.

Представленные в списке Top500 данные позволяют проследить характерные тенденции развития индустрии в сфере суперкомпьютерных вычислений. Первый список Top500 датирован июнем 1993 г. и содержал 249 многопроцессорных систем с общей памятью и 97 суперкомпьютеров, построенных на основе единственного процессора; более 40% всех решений в нем были созданы на платформе, разработанной компанией Cray. Уже четыре-мя годами позже в Top500 не осталось ни одного суперкомпьютера на основе единственного процессора, а взамен появилась первая система с производительностью всего в 10 гигафлопс (в 100 раз меньше, чем у лидера списка системы ASCI Red), относящаяся к довольно новому тогда виду кластерных вычислительных систем, которые сегодня занимают в Top500 более

---

<sup>4</sup> Значительно повышаются энергопотребление процессора и его системы охлаждения.

80% списка и являются, фактически, основным способом построения суперкомпьютеров.

Основным преимуществом кластеров, предопределившим их повсеместное распространение, было и остается построение из стандартных массово выпускающихся компонентов, как аппаратных, так и программных. В 31-м списке Top500 75% систем построены на основе процессоров компании Intel, чуть больше 13% – на процессорах компании IBM и 11% – компании AMD (на двух оставшихся производителей – NEC и Cray – приходится по одной системе соответственно); 81% систем используют всего два типа сетей передачи данных: Gigabit Ethernet или Infiniband; 85% систем работают под управлением операционной системы из семейства Linux. Как видим, список использующихся программно-аппаратных компонент весьма ограничен, что является несомненным плюсом с точки зрения пользователей.

Однако для массового пользователя еще большим плюсом была бы возможность иметь персональный суперкомпьютер у себя на столе или в «облаке» с надёжным и быстрым доступом к нему. И кластеры, принесшие в индустрию высокопроизводительных вычислений идею «собери суперкомпьютер своими руками», как нельзя лучше отвечают этой потребности. Сейчас трудно достоверно установить, какая система может быть названа первым в мире «персональным кластером». Во всяком случае, уже в начале 2001 г. компания RenderCube [109] представила одноименный мини-кластер из 4-х двухпроцессорных систем, заключенных в кубический корпус со стороной всего в 42 см.

Тенденция «персонализации» супервычислений в последнее время развивается все активнее, и недавно была подхвачена в том числе и производителями видеокарт, мощности которых возросли настолько, что возникло естественное желание использовать их не только в графических расчетах, но и в качестве ускорителей вычислений общего назначения. Соответствующие решения представлены в настоящее время компанией NVIDIA (семейство NVIDIA® Tesla™) и компанией AMD (семейство ATI FireStream™) и демонстрируют – в силу специфики внутреннего устройства – потрясающую (в сравнении с универсальными процессорами) пиковую производительность, превышающую 1 терафлопс.

Основная идея кластера в "облаке" заключается в предоставлении пользователям вычислительных ресурсов кластера удалённо через сеть Internet. При этом, пользователю не нужно покупать кластер, заниматься его содержанием и обслуживанием. Кроме того, оплачиваются только потреблённые ресурсы при выполнении вычисления. Одной из популярных облачных систем является Amazon Web Services (в 41-ом списке Top500 кластер Amazon EC2 занимает 128 место с производительностью 240 терафлопс).

Данная глава посвящена рассмотрению современных многоядерных процессоров, которые являются основой для построения самых быстрых дей-

ствующих вычислительных систем. Для полноты картины приводится также описание ряда аппаратных устройств (видеокарт и вычислительных сопроцессоров), которые могут быть использованы для существенного ускорения вычислений.

## 1. Параллелизм как основа высокопроизводительных вычислений

Без каких-либо особых преувеличений можно заявить, что все развитие компьютерных систем происходило и происходит под девизом «Скорость и быстрота вычислений». Если быстроедействие первой вычислительной машины ENIAC составляло всего несколько тысяч операций в секунду, то самый быстрый на ноябрь 2012 г. суперкомпьютер Titan может выполнять уже несколько квадриллионов ( $10^{15}$ ) команд. Темп развития вычислительной техники просто впечатляет – увеличение скорости вычислений в триллионы ( $10^{12}$ ) раз немногим более чем за 60 лет! Для лучшего понимания необычности столь стремительного развития средств вычислительной техники часто приводят яркие аналогии, например: если бы автомобильная промышленность развивалась с такой же динамикой, то сейчас автомобили весили бы порядка 200 граммов и тратили бы несколько литров бензина на миллионы километров!

История развития вычислительной техники представляет увлекательное описание замечательных научно-технических решений, радости побед и горечи поражений. Проблема создания высокопроизводительных вычислительных систем относится к числу наиболее сложных научно-технических задач современности, и ее разрешение возможно только при всемерной концентрации усилий многих талантливых ученых и конструкторов, предполагает использование всех последних достижений науки и техники и требует значительных финансовых инвестиций. Здесь важно отметить, что при общем росте скорости вычислений в  $10^{12}$  раз быстроедействие самих технических средств вычислений увеличилось всего в несколько миллионов раз. И дополнительный эффект достигнут за счет введения параллелизма буквально на всех стадиях и этапах вычислений.

Не ставя целью в рамках данного курса подробное рассмотрение истории развития компьютерного параллелизма, отметим, например, организацию независимости работы разных устройств ЭВМ (процессора и устройств ввода-вывода), появление многоуровневой памяти, совершенствование архитектуры процессоров (суперскалярность, конвейерность, динамическое планирование). Дополнительная информация по истории параллелизма может быть получена, например, в [67]; здесь же выделим как принципиально важный итог – многие возможные пути совершенствования процес-

соров практически исчерпаны (так, возможность дальнейшего повышения тактовой частоты процессоров ограничивается рядом сложных технических проблем) и наиболее перспективное направление на данный момент времени состоит в явной организации многопроцессорности вычислительных устройств.

Ниже будут более подробно рассмотрены основные способы организации многопроцессорности – симметричной мультипроцессорности (Symmetric Multiprocessor, SMP), одновременной многопоточности (Simultaneous Multithreading, SMT) и многоядерности (multicore).

### 1.1. Симметричная мультипроцессорность

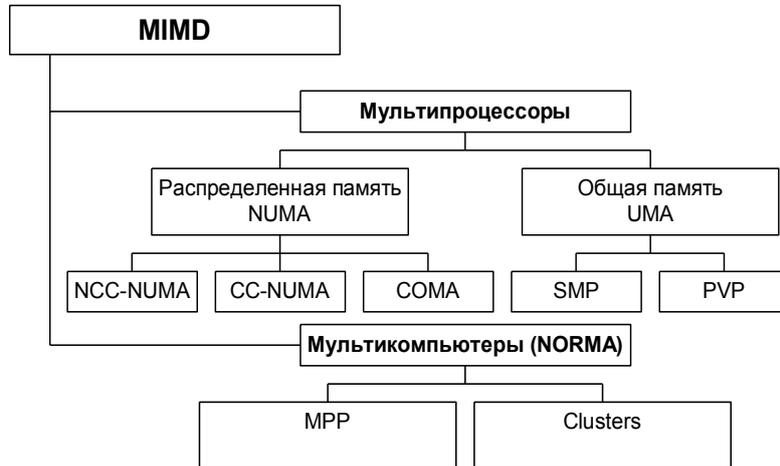
Организация симметричной мультипроцессорности (Symmetric Multiprocessor, SMP), когда в рамках одного вычислительного устройства имеется несколько полностью равноправных процессоров, является практически первым использованным подходом для обеспечения многопроцессорности – первые вычислительные системы такого типа стали появляться в середине 50-х – начале 60-х годов, однако массовое применение SMP систем началось только в середине 90-х годов.

Следует отметить, что SMP системы входят в группу MIMD (multi instruction multi data) вычислительных систем в соответствии с классификацией Флинна. Поскольку эта классификация приводит к тому, что практически все виды параллельных систем (несмотря на их существенную разнородность) относятся к одной группе MIMD, для дальнейшей детализации класса MIMD предложена практически общепризнанная структурная схема [47, 99] – см. рис. 1. В рамках данной схемы дальнейшее разделение типов многопроцессорных систем основывается на используемых способах организации оперативной памяти в этих системах. Данный подход позволяет различать два важных типа многопроцессорных систем – multiprocessors (мультипроцессоры, или системы с общей разделяемой памятью) и multicomputers (мультикомпьютеры, или системы с распределенной памятью).

Для дальнейшей систематики мультипроцессоров учитывается способ построения общей памяти. Возможный подход – использование единой (централизованной) общей памяти (shared memory) – см. рис. 2. Такой подход обеспечивает однородный доступ к памяти (uniform memory access, UMA) и служит основой для построения векторных параллельных процессоров (parallel vector processor, PVP) и симметричных мультипроцессоров (symmetric multiprocessor, SMP). Среди примеров первой группы – суперкомпьютер Cray T90, ко второй группе относятся IBM Server, Sun StarFire, HP Superdome, SGI Origin и др.

Одной из основных проблем, которые возникают при организации параллельных вычислений на такого типа системах, является обеспечение ин-

формационной целостности (когерентности) кэшей (cache coherence problem).



**Рис. 1.** Классификация многопроцессорных вычислительных систем

Дело в том, что при наличии общих данных, копии значений одних и тех же переменных могут оказаться в кэшах разных процессоров. Если в такой ситуации (при наличии копий общих данных) один из процессоров выполнит изменение значения разделяемой переменной, то значения копий в кэшах других процессорах окажутся несоответствующими действительности и их использование приведет к некорректности вычислений.

Существует две методики дублирования данных их кэша в оперативную память: обратная запись (write back) и сквозная запись (write through). При использовании сквозной записи в кэш все записываемые данные сразу дублируются в оперативной памяти. При использовании обратной записи данные помещаются только в кэш и переписываются в оперативную память только тогда, когда необходимо освободить строку кэша. При использовании этих методик может нарушиться информационная целостность данных. Так, при использовании обратной записи модификация строки кэша никак не отразится на состоянии копий этой строки в других кэшах и оперативной памяти. Остальные процессоры не будут «знать», что в их кэшах находятся устаревшие данные. Для обеспечения информационной целостности данных при использовании сквозной записи необходимо дублировать записываемые данные не только в оперативную память, но и во все кэши, которые содержат эти данные, что негативно сказывается на эффективности системы.

Обеспечение однозначности кэшей может быть достигнуто на программном или аппаратном уровнях. Вся тяжесть обеспечения информационной целостности кэшей при использовании программных методов ложится на

операционную систему и компилятор. При сборке программы компилятор должен определить те переменные и элементы данных, которые могут потенциально нарушить целостность данных и запретить их кэширование. Реализовать подобную стратегию эффективно очень тяжело, т.к. запрещать кэширование участков памяти необходимо только в тех случаях, когда это требуется для обеспечения однозначности. Блокирование кэширования во всех остальных случаях (когда информационной целостности данных ничто не угрожает) приведёт к необоснованному снижению быстродействия программы.

Обеспечение однозначности кэшей обычно реализуется на аппаратном уровне. Аппаратные методы позволяют выполнять действия, необходимые для обеспечения когерентности данных, только в тех случаях, когда это требуется при выполнении программы, что приводит к более эффективному использованию кэшей. Существует множество аппаратных методов обеспечения информационной целостности данных. Более подробно об этом можно прочитать в [1]. Мы рассмотрим метод с использованием протокола MESI.

Протокол MESI основан на введении четырёх состояний для каждой строки кэша:

*modified* (строка была изменена и изменения не отражены в оперативной памяти),

*exclusive* (данные в строке кэша и оперативной памяти одинаковы, в остальных кэшах этих данных нет),

*shared* (данные в строке кэша и оперативной памяти одинаковы, в каких-то других кэшах этих данные тоже присутствуют),

*invalid* (строка содержит недостоверные данные).

При выполнении операций чтения и записи может возникнуть одна из двух ситуаций: промах (данных нет в кэше, либо данные недостоверны) или попадание (данные присутствуют в кэше). Операция чтения при попадании в кэш никак не изменяет состояния строки. Операция записи при попадании выполняется по-разному, в зависимости от состояния строки. Если строка имела состояние *exclusive* или *modified*, то текущий процессор имеет исключительное право владения этой строкой и изменяет состояние строки на *modified*. Если строка имела состояние *shared*, то состояние копии строки в других кэшах меняется на *invalid*, а текущий процессор обновляет данные в кэше и изменяет состояние строки на *modified*. Операции чтения при промахе может выполняться по-разному в зависимости от состояния других кэшей. Если в каком-либо кэше содержится копия строки в состоянии *exclusive* или *shared*, то данные загружаются в кэш из оперативной памяти, а строка помечается как *shared* в этих кэшах. Если в каком-либо кэше содержится копия строки в состоянии *modified*, то этот кэш блокирует чтение из памяти и иницирует запись строки в оперативную память,

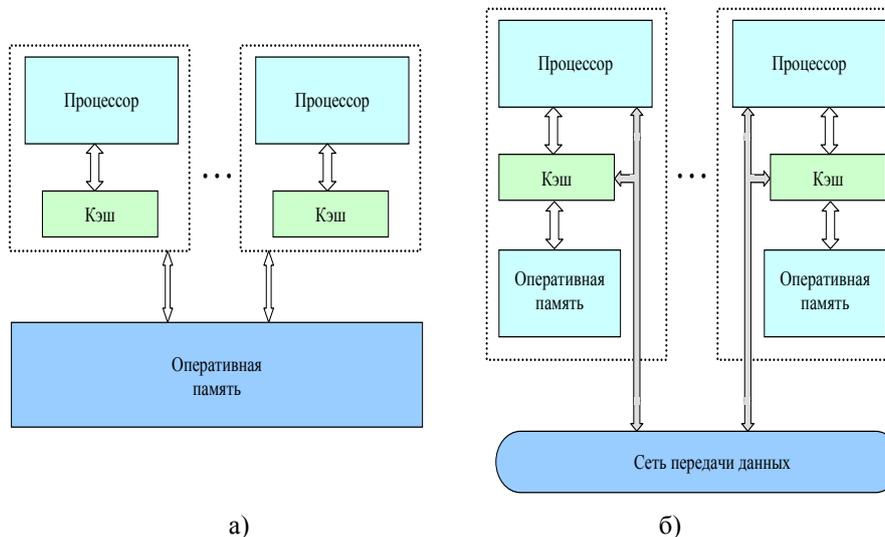
после чего изменит состояние строки на shared. Если ни в одном из кэшей нет копии строки, то блок данных загружается из оперативной памяти, а строка помечается как exclusive. Операция записи при промахе инициирует чтение блока данных из оперативной памяти и устанавливает состояние строки в значение modified. При этом, если в каком-либо кэше находилась копия данных в состоянии modified, то сначала данные из этого кэша записываются в оперативную память, а строка помечается как invalid. Все кэши, которые содержали копию данных в состоянии shared или exclusive переходят в состояние invalid.

Рассмотренный выше механизм будет работать только для кэшей, которые имеют выход на общую магистраль данных. По магистрали данных циркулируют потоки информации между отдельными процессорами и оперативной памятью и сигналы, необходимые для обеспечения когерентности данных к кэшам. При увеличении количества процессоров, поток данных по магистрали тоже увеличивается. Это приводит к снижению скорости вычислений и затрудняет создание систем с достаточно большим количеством процессоров. Как правило, на общую магистраль имеют выходы кэши L2. Для обеспечения когерентности кэшей L1 используются более сложные алгоритмы (особенно при использовании обратной записи).

На данный момент широко используется модифицированный протокол MESI в котором добавлено дополнительное состояние, позволяющее эффективнее решать проблему когерентности данных за счёт загрузки требуемых данных из соседнего кэша, а не из оперативной памяти. Intel использует протокол MESIF, AMD – MOESI.

Наличие общих данных при выполнении параллельных вычислений приводит к необходимости синхронизации взаимодействия одновременно выполняемых потоков команд. Так, если изменение общих данных требует для своего выполнения некоторой последовательности действий, то необходимо обеспечить взаимное исключение (mutual exclusion), с тем чтобы эти изменения в любой момент времени мог выполнять только один командный поток. Задачи взаимного исключения и синхронизации относятся к числу классических проблем, и их рассмотрение при разработке параллельных программ является одним из основных вопросов параллельного программирования.

Общий доступ к данным может быть обеспечен и при физически распределенной памяти (при этом, естественно, длительность доступа уже не будет одинаковой для всех элементов памяти) – см. рис. 2. Такой подход именуется как неоднородный доступ к памяти (non-uniform memory access, или NUMA). Среди систем с таким типом памяти выделяют:



а) б)  
**Рис. 2.** Архитектура многопроцессорных систем с общей (разделяемой) памятью: системы с (а) однородным и (б) неоднородным доступом к памяти

- Системы, в которых для представления данных используется только локальная кэш-память имеющихся процессоров (*cache-only memory architecture*, или *COMA*); примерами таких систем являются KSR-1 и DDM.
- Системы, в которых обеспечивается когерентность локальных кэшей разных процессоров (*cache-coherent NUMA*, или *CC-NUMA*); среди систем данного типа – SGI Origin 2000, Sun HPC 10000, IBM/Sequent NUMA-Q 2000.
- Системы, в которых обеспечивается общий доступ к локальной памяти разных процессоров без поддержки на аппаратном уровне когерентности кэша (*non-cache coherent NUMA*, или *NCC-NUMA*); к данному типу относится, например, система Cray T3E.

Использование распределенной общей памяти (*distributed shared memory*, *DSM*) упрощает проблемы создания мультипроцессоров (известны примеры систем с несколькими тысячами процессоров), однако возникающие при этом проблемы эффективного использования распределенной памяти (время доступа к локальной и удаленной памяти может различаться на несколько порядков) приводят к существенному повышению сложности параллельного программирования.

## 1.2. Одновременная многопоточность

Организация симметричной мультипроцессорности позволяет достаточно легко увеличивать производительность вычислительных устройств. Однако такое решение при увеличении числа процессоров обладает плохой

масштабируемостью из-за проблем с обеспечением когерентности кэш-памяти разных процессоров – SMP системы содержат, как правило, 2 или 4, реже – 8 и совсем редко – большее количество процессоров. Кроме того, такой подход является сравнительно дорогим решением за счёт дорогой реализации компонент (требуются сложные и дорогие материнские платы для размещения нескольких процессоров).

С другой стороны, проанализировав эффективность современных сложных процессоров, насчитывающих в своем составе десятки и сотни миллионов транзисторов, можно обратить внимание на то, что при выполнении большинства операций оказываются полностью задействованными не все составные компоненты процессоров (по имеющимся оценкам, средняя загрузка процессора составляет всего лишь около 30%). Так, если в данный момент времени выполняется операция целочисленной арифметики, то блок процессора для выполнения вещественных операций окажется простаивающим. Для повышения загрузки процессора можно организовать спекулятивное (опережающее) исполнение операций, однако воплощение такого подхода требует существенного усложнения логики аппаратной реализации процессора. Было бы гораздо проще, если бы в программе заранее были предусмотрены последовательности команд (потoki), которые могли быть выполнены параллельно и независимо друг от друга. Такой подход тем более является целесообразным, поскольку поддержка многопоточного исполнения может быть обеспечена и на аппаратном уровне за счет соответствующего расширения возможностей процессора (и такая доработка является сравнительно простой).

Данная идея поддержки одновременной многопоточности (simultaneous multithreading, SMT) была предложена в 1995 г. в университете Вашингтона Динем Тулсенем (Dean Tullsen) и позднее активно развита компанией Intel под названием технологии гиперпоточности (hyper threading, HT). В рамках такого подхода процессор дополняется средствами запоминания состояния потоков, схемами контроля одновременного выполнения нескольких потоков и т. д. За счет этих дополнительных средств на активной стадии выполнения может находиться несколько потоков; при этом одновременно выполняемые потоки конкурируют за исполнительные блоки единственного процессора, и как результат выполнение отдельных потоков может блокироваться, если требуемые в данный момент времени блоки процессора оказываются уже задействованными. Как правило, число аппаратно-поддерживаемых потоков равно 2, в более редких случаях этот показатель достигает 4 и даже 8. Важно при этом подчеркнуть, что аппаратно-поддерживаемые потоки на логическом уровне операционных систем Linux и Windows воспринимаются как отдельные процессоры, т. е., например, единственный процессор с двумя аппаратно-поддерживаемыми потоками в менеджере Task Manager операционной системы Windows диагностируется как два отдельных процессора.

Использование процессоров с поддержкой многопоточности может приводить к существенному ускорению вычислений (важно отметить – при надлежащей реализации программ). Так, имеется большое количество демонстраций, показывающих, что на процессорах компании Intel с поддержкой технологии гиперпоточности достигается повышение скорости вычислений около 30%.

Стоит заметить, что одновременная многопоточность не всегда положительно сказывается на производительности. При использовании многопоточности несколько программных потоков делят ресурсы одного вычислительного ядра, поэтому для высокопроизводительных приложений (где потоки выполняют однотипные вычисления и конкурируют за ресурсы ядра) эта технология часто сказывается негативно. Выигрыш от использования одновременной многопоточности может достигаться, если на одном ядре выполняются потоки разнотипных приложений (например, просмотр почты и проигрывание музыки), если на одном ядре выполняется несколько потоков с интенсивным вводом-выводом и пр.

### 1.3. Многоядерность

Технология одновременной многопоточности позволяет достичь многопроцессорности на логическом уровне. Еще раз отметим, что затраты на поддержку такой технологии являются сравнительно небольшими, но и получаемый результат достаточно далек от максимально-возможного – ускорение вычислений от использования многопоточности оказывается равным примерно 30%. Дальнейшее повышение быстродействия вычислений при таком подходе по-прежнему лежит на пути совершенствования процессора, что – как было отмечено ранее – требует решения ряда сложных технологических проблем.

Возможное продвижение по направлению к большей вычислительной производительности может быть обеспечено за счёт реализации в единственном кремниевом кристалле несколько вычислительных ядер в составе одного многоядерного процессора, при этом по своим вычислительным возможностям эти ядра не уступают обычным процессорам.

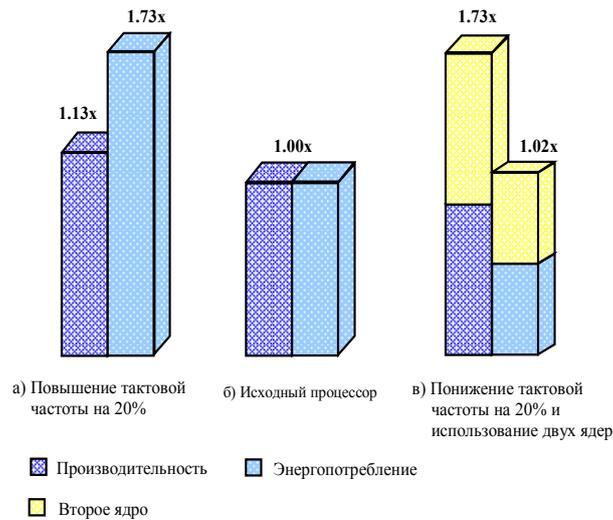
Кроме того, можно пойти по пути возврата к более «простым» процессорам с более низкой тактовой частотой и с менее сложной логикой реализации! Такой неординарный ход приводит к тому, что процессоры становятся менее энергоёмкими<sup>5</sup>, более простыми для изготовления и как результат – более надёжными. А также – что является чрезвычайно важным – «простые» процессоры требуют для своего изготовления меньшее количество логических схем, что приводит к освобождению – в рамках кремниевого

---

<sup>5</sup> Проблема энергопотребления является одной из наиболее сложных для процессоров с высокой тактовой частотой.

кристалла, используемого для изготовления процессоров, – большого количества свободных транзисторов. Эти свободные транзисторы, в свою очередь, могут быть использованы для реализации дополнительных вычислительных устройств, которые могут быть добавлены к процессору<sup>6</sup>.

Поясним сказанное на примере рис. 3. В центре рисунка (рис. 3б) приведены показатели исходного процессора, принятые за 1 для последующего сравнения.



**Рис. 3.** Демонстрация зависимости между тактовой частотой, энергопотреблением и производительностью процессора

Пусть для повышения быстродействия процессора его тактовая частота увеличена на 20% (см. рис. 3а), тогда производительность процессора увеличится – однако не на 20%, а, например, на 13% (приводимые здесь числовые значения имеют качественный характер), в то время как энергопотребление возрастет существенно – в приведенном примере на 73%. Данный пример является на самом деле очень характерным – увеличение тактовой частоты процессора приводит в большинстве случаев к значительному росту энергопотребления. Теперь уменьшим тактовую частоту процессора – опять же на 20% (см. рис. 3в). В результате снижения тактовой частоты производительность процессора, конечно, уменьшится, но опять же не на 20%, а примерно на ту же величину 13% (т. е. станет равной 87% от производительности исходного процессора). И опять же, энергопотребление процессора уменьшится, причем достаточно значительно – до уров-

<sup>6</sup> Подобный подход реализован во многих сопроцессорах, часть из которых будет рассмотрена далее.

ня примерно 51% энергопотребления исходного процессора. И тогда, добавив в процессор второе вычислительное ядро за счет появившихся свободных транзисторов, мы можем довести суммарные показатели процессора по энергопотреблению до уровня 1.02 энергопотребления исходного процессора, а производительность – до уровня 1.73!!!

На логическом уровне архитектура многоядерного процессора соответствует практически архитектуре симметричного мультипроцессора (рис. 2 и 4). На рис. 4 приведена возможная архитектура двухядерного процессора – различия для разных многоядерных процессоров могут состоять в количестве имеющихся ядер и в способах использования кэш-памяти ядрами процессора – кэш-память может быть как общей, так и распределенной для разных ядер. Так, на рис. 4 кэш-память первого уровня L1 локальна для каждого ядра, в то же время кэш-память всех последующих уровней и оперативная память является общей.

Как следует из проведенного рассмотрения, многоядерность позволяет повышать производительность процессоров, и данный подход обладает целым рядом привлекательных моментов (уменьшение энергопотребления, снижение сложности логики процессоров и т. п.). Все сказанное приводит к тому, что многоядерность становится одним из основных направлений развития компьютерной техники.

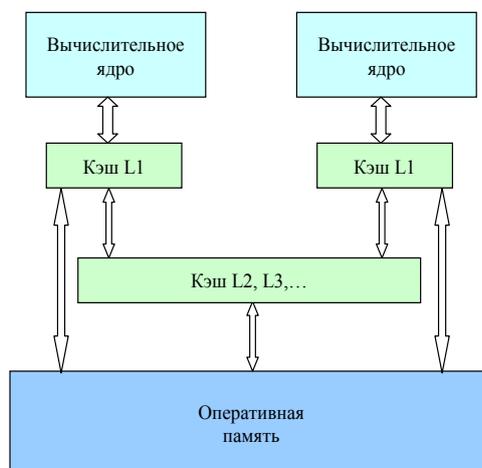


Рис. 4. Архитектура двухядерного процессора

Значимость такого подхода привела даже к тому, что известный закон Мура теперь формулируется в виде «Количество вычислительных ядер в процессоре будет удваиваться каждые 18 месяцев». Стоит отметить, что удвоение числа ядер не приводит к мгновенному увеличению производительности, т.к. требуется время для разработки алгоритмов способных эффективно использовать эти ядра. В настоящее время для массового использования доступны двух-, четырех-, шести- и восьми- ядерные процессоры. В научно-технической литературе наряду с рассмотрением обычных мно-

гоядерных (multi-core) процессоров начато широкое обсуждение процессоров с массовой многоядерностью (many-core), когда в составе процессоров будут находиться сотни и тысячи ядер!

И в заключение следует отметить еще один принципиальный момент – потенциал производительности многоядерных процессоров может быть задействован только при надлежащей разработке программного обеспечения – программы должны быть очень хорошо распараллелены. А, как известно, сложность разработки параллельных программ значительно превышает трудоемкость обычного последовательного программирования. Тем самым проблема обеспечения высокопроизводительных вычислений перемещается теперь из области компьютерного оборудования в сферу параллельного программирования. И здесь нужны новые идеи и перспективные технологии для организации массового производства параллельных программ.

## **2. Многоядерность – два, четыре, восемь – кто больше?**

Дополним теперь общее рассмотрение многоядерного направления развития компьютерной техники характеристикой ряда конкретных широко применяемых в настоящее время многоядерных процессоров основных компаний-разработчиков – Intel, AMD, IBM, Sun (Oracle).

### **2.1. Процессоры Intel® Core™ и Intel® Xeon®**

Как просто было когда-то сравнивать процессоры компании Intel между собой. Все знали: есть Pentium, есть его «урезанный» вариант Celeron, а в остальном чем выше частота, тем лучше. Эта простота была следствием того факта, что в формуле, определяющей производительность вычислительной системы, – «тактовая частота процессора × число инструкций, выполняемых за один такт (Instructions Per Cycle, IPC)», – переменной величиной была только частота. Необходимо, конечно, отметить, что получаемая по этой формуле величина дает только так называемую «пиковую производительность», приблизиться к которой на практике можно лишь на отдельных специально подобранных задачах. Именно поэтому сравнение вычислительных систем, в том числе из списка Top500 и аналогичных им, выполняется на основе производительности, показанной на стандартном тесте, в качестве которого повсеместно используется LINPACK [110]. Как только наращивание тактовой частоты прекратилось, компании Intel понадобился другой способ описания и градации выпускаемых процессоров. Сегодня все процессоры производства Intel делятся прежде всего по назначению: для настольных систем, ноутбуков, серверов и рабочих станций и т.д. Затем в каждом классе выделяют серии процессо-

ров, отличающиеся между собой по некоторым ключевым характеристикам.

В сентябре 2006 г. Intel анонсировала новую стратегию развития микропроцессоров “Tick-Tock”, состоящую из двух стадий. На первой стадии “Tick” происходит уменьшение техпроцесса и незначительное усовершенствование архитектуры. На стадии “Tock” происходит выпуск новой микроархитектуры на существующем техпроцессе.

## The Tick-Tock model through the years

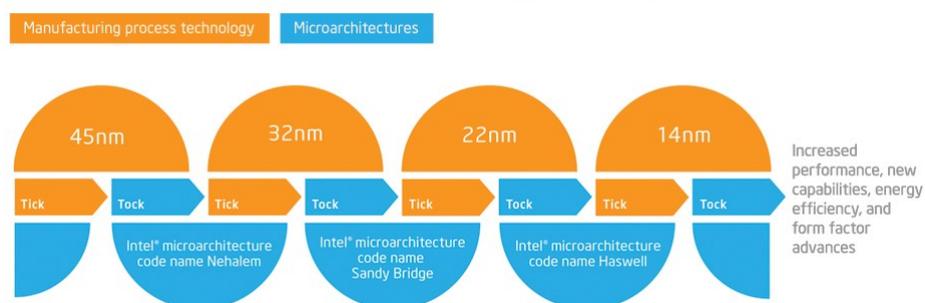


Рис. 5. Стратегия “Tick-Tock”<sup>7</sup>

На данный момент широко распространены процессоры Intel с микроархитектурой Sandy Bridge, основанной на 32-нм технологии, и Ivy Bridge, которая является развитием предыдущей и основана на 22-нм технологии с использованием транзисторов с трехмерной структурой затвора (3D Tri-Gate Transistor). В классе настольных и мобильных систем сегодня «царствуют» представители семейства Intel® Core™ 3-го поколения, в серверном сегменте – процессоры Intel® Xeon®.

Особенностям микроархитектуры Intel® Core™, позволившей компании Intel существенно потеснить своего основного конкурента – компанию AMD, посвящено множество материалов и публикаций. Не ставя перед собой задачу подробного ее обсуждения, кратко отметим лишь ключевые моменты, выделяемые самими разработчиками.

- В процессор интегрирован северный мост набора системной логики (2-канальный контроллер памяти DDR3 и контроллер шины PCI-Express 3.0).
- Smart Cache. Динамически распределяет кэш-память 3-го уровня между ядрами процессора в зависимости от нагрузки.

<sup>7</sup> <http://www.intel.com/content/www/us/en/silicon-innovations/intel-tick-tock-model-general.html>

- Turbo Boost 2.0. В случае загрузки не всех процессорных ядер, эта технология обеспечивает автоматическое увеличение тактовой частоты работы тех, которые остаются активными.
- Hyper-Threading (см. одновременная многопоточность). Эта технология позволяет удвоить число ядер, доступных операционной системе, за счёт обработки нескольких потоков команд на каждом ядре. У четырёхъядерных процессоров с технологией Hyper-Threading операционной системе будет доступно восемь ядер.
- Advanced Vector Extension (AVX) Technology. Набор новых команд для векторных вычислений (каждый регистр имеет размер 256 бит). Команды предназначены для повышения производительности приложений с большими объемами вычислений с плавающей запятой.
- Встроенная графическая система Intel HD Graphics.
- Advanced Encryption Standard New Instructions (AES-NI). Набор инструкций процессора, которые обеспечивают быстрое и надежное шифрование данных по стандарту AES.

Приведем технические данные текущих лидеров в классе настольных и серверных процессоров (оба процессора поддерживают технологию Hyper-Threading).

#### **Процессор Intel® Core™ i7-3770**

Тактовая частота: 3,4 ГГц (максимальная 3,9 ГГц).

Число ядер: 4.

Кэш второго уровня: 1 МБ (по 256 КБ на каждое ядро).

Кэш третьего уровня: 8 МБ.

Технологический процесс: 22 нанометров.

#### **Процессор Intel® Xeon® Processor E7-8830**

Тактовая частота: 2,13 ГГц (максимальная 2,4 ГГц).

Число ядер: 8.

Кэш второго уровня: 16 МБ (по 2 МБ на каждое ядро).

Кэш третьего уровня: 24 МБ.

Технологический процесс: 32 нанометров.

В заключение отметим еще один весьма важный факт – помимо пиковой производительности той или иной архитектуры и соответственно процессоров, построенных на ее основе, значимым обстоятельством для конечного потребителя является процент мощности, который можно получить от пика. Для систем в Top500, построенных на процессорах компании Intel, этот показатель составляет в 40-м списке 64,7%, при этом «удельная мощность» в расчете на один процессор/ядро равна 16,5 гигафлопс (необходи-

мо заметить, что значительная часть этих систем введена в строй уже несколько лет назад и построена не на новейших процессорах, а часть систем содержит ускорители и сопроцессоры, что значительно повышает их производительность).

## 2.2. Процессоры AMD Phenom™ и AMD Opteron™

Компания AMD основана в 1969 г. (всего на год позже, чем Intel) и в сознании рядового пользователя прочно занимает место главного конкурента Intel на рынке процессоров для настольных систем и отчасти на рынке серверных, при этом практически всегда выступая в роли догоняющего. Если принимать во внимание только «внешние» факторы вроде рыночной доли, то ситуация действительно может быть воспринята именно так. И в этом свете основной успех компании за последнее десятилетие связан с выпуском в 2003 г. 64-битных процессоров AMD Opteron™, быстро завоевавших популярность и позволивших AMD значительно упрочить свое положение, в том числе в сегменте высокопроизводительных решений. Достаточно отметить, что в 28-м списке Top500 (ноябрь 2006 г.) доля систем, построенных на основе процессоров AMD, достигла своего исторического максимума и составила 22,6% против 52,6% у компании Intel и 18% у компании IBM. Однако кроме такого чисто количественного сравнения, в котором AMD неизменно проигрывает своим конкурентам, есть еще показатели качественные, и тут компания за прошедшие годы нередко бывала первопроходцем и реализовывала действительно интересные архитектурные решения.

Среди прочего это и интеграция в процессор северного моста, что дает более быстрый доступ к оперативной памяти, и использование Direct Connect Architecture для взаимодействия процессоров между собой посредством высокоскоростной шины HyperTransport™, позволяющей без существенных потерь в производительности объединять в рамках одной системы до 8 процессоров Opteron. Стоит отметить, что процессоры AMD поддерживают технологии, аналогичные рассмотренных ранее у Intel: расширенная технология оптимизации энергопотребления (AMD PowerNow!™ Technology), увеличение частоты ядер при необходимости (AMD Turbo CORE Technology), поддержка новых инструкций (AVX, AES, FMA4, XOP), общий кэш третьего уровня (AMD Balanced Smart Cache) и многое другое. Последними серверными процессорами компании AMD являются шестнадцатиядерные модели Opteron на ядре Piledriver.

На рынке настольных систем основное оружие компании AMD сегодня – процессоры AMD FX™. Процессоры FX построены на той же микроархитектуре (Piledriver), что и серверные Opteron. Текущее поколение процессоров FX выпускается по технологии 32 нм.

Приведем технические данные текущих лидеров в классе настольных и серверных процессоров.

**Процессор AMD FX 8350**

Тактовая частота: 4,0 ГГц (максимальная 4,2 ГГц).

Число ядер: 8.

Кэш второго уровня: 8 МБ (по 1 МБ на каждое ядро).

Кэш третьего уровня: 8 МБ (общий на все ядра).

Технологический процесс: 32 нанометров.

**Процессор AMD Opteron™ 6386 SE**

Тактовая частота: 2,8 ГГц (максимальная 3,5 ГГц).

Число ядер: 16.

Кэш второго уровня: 16 МБ (по 1 МБ на каждое ядро).

Кэш третьего уровня: 16 МБ (общий на все ядра).

Технологический процесс: 32 нанометров.

В заключение, как и для процессоров компании Intel, приведем усредненные данные из списка Top500. Для систем в Top500, построенных на процессорах компании AMD, отношение «показанная мощность/пиковая мощность» составляет в 40-м списке 70,2%, при этом «удельная мощность» в расчете на один процессор/ядро равна 17,85 гигафлопс. Как и ранее, отметим, что значительная часть этих систем введена в строй уже несколько лет назад и построена не на новейших процессорах, а часть систем содержит ускорители и сопроцессоры, что значительно повышает их производительность. Отметим, что лидер 40-го списка TOP-500 Titan построен на 16-ядерных процессорах AMD Opteron 6274.

**2.3. Процессоры IBM Power7**

История компании IBM значительно длиннее, чем у Intel и AMD, и, в отличие от последних, IBM никогда не производила только процессоры. Фактически компания, говоря сегодняшним языком, всегда пыталась поставлять «готовые решения». Однако обсуждения всего списка продукции IBM выходит за рамки данного материала, и мы остановимся только на процессорах, которые выпускает компания сегодня и на основе которых строит как серверы «начального» уровня, так и суперкомпьютеры вроде Roadrunner или BlueGene.

Микропроцессорная архитектура Power (расшифровывается как Performance Optimization With Enhanced RISC) имеет не менее богатую историю, чем сама компания IBM. Начиная с 1990 г., когда были выпущены первые компьютеры на основе процессоров Power, и по сегодняшний день архитектура постоянно развивается, с каждым поколением процессоров при-

внося значительные новшества. Текущая версия процессоров Power – Power7 выпущена в 2010 г., тем не менее уже 13 систем в 40-м списке Top500 построено на основе этих процессоров.

Процессор Power7 выпускается по 45 нм технологическому процессу. Максимальная частота серийно выпускаемых образцов на сегодня равна 4,25 ГГц.

Процессоры Power7 могут иметь четыре, шесть или восемь ядер, способных выполнять по четыре потока команд одновременно, по 4 МБ кэша третьего уровня на каждое ядро. Каждое ядро содержит два блока работы с целыми числами и четыре – с числами с плавающей точкой. Процессор поддерживает внеочередное (out-of-order<sup>8</sup>) исполнения команд и каждое ядро способно выполнять до 8 инструкций за такт. Процессоры Power7 поставляются в многочиповом корпусе, вмещающем до 4 процессоров.

В заключение, как и ранее, приведем усредненные данные из списка Top500. Для систем в Top500, построенных на процессорах IBM семейства Power, отношение «показанная мощность/пиковая мощность» составляет в 40-м списке 81%, при этом «удельная мощность» в расчете на один процессор/ядро равна 12,5 гигафлопс. Как и ранее, отметим, что значительная часть этих систем построена не на новейших процессорах.

В августе 2012 IBM представила процессоры Power7+, которые являются развитием процессоров Power7 с большей тактовой частотой (4,4 ГГц) и размером кэша (10 МБ на ядро). Процессоры Power7+ выпускается по 32 нм технологическому процессу.

#### 2.4. Процессоры PowerXCell™ 8i

Рассказ о процессоре PowerXCell™ 8i конечно же нужно начать с его прямого предка – процессора Cell (рис. 6), разработанного альянсом STI (Sony, Toshiba, IBM) в первую очередь для использования в игровых приставках Sony PlayStation 3. В процессе создания этого процессора были приняты весьма интересные решения, дающие в итоге очень высокую пиковую производительность (более 200 гигафлопс, правда, только для вещественной арифметики одинарной точности), но требующие в качестве платы более сложного программирования.

---

<sup>8</sup> Процессор может изменить порядок выполнения инструкций (если это возможно), чтобы не допустить простаивания своих вычислительных блоков. Внеочередное выполнение команд поддерживается многими современными процессорами. В том числе процессорами Intel и AMD.

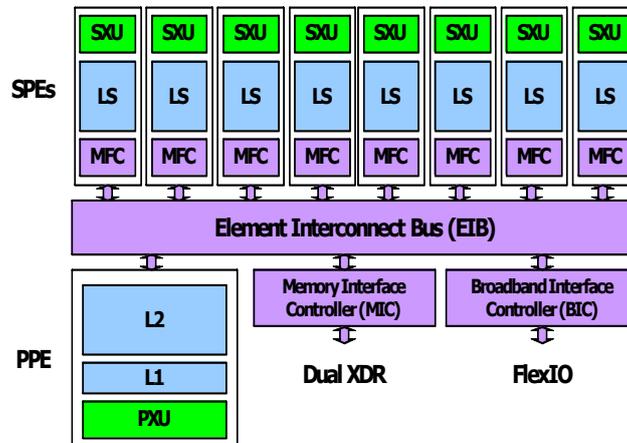


Рис. 6. Процессор Cell

Прежде всего отметим, что процессор Cell имеет существенно «неоднородное» устройство. Он состоит из одного двухъядерного Power Processor Element (PPE) и 8 Synergistic Processor Element (SPE). PPE построен на архитектуре PowerPC и «отвечает» в процессоре Cell за исполнение кода общего назначения (операционной системы в частности), а также контролирует работу потоков на сопроцессорах SPE. Ядра PPE – 64-разрядные и так же, как и Power6, используют поочередный (in-order) порядок исполнения команд. PPE имеет блок векторных операций Vector Multimedia eXtensions (VMX), кэш первого уровня размеров 64 КБ (по 32 КБ на кэш инструкций и данных) и кэш второго уровня размером 512 КБ.

В отличие от PPE, SPE-ядра представляют собой специализированные векторные процессоры, ориентированные на быструю потоковую работу с SIMD-инструкциями. Архитектура SPE довольно проста: четыре блока для работы с целочисленными векторными операциями и четыре блока для работы с числами с плавающей запятой. Большинство арифметических инструкций представляют данные в виде 128-разрядных векторов, разделенных на четыре 32-битных элемента. Каждый SPE оснащён 128 регистрами, разрядность которых – 128-бит. Вместо кэша первого уровня SPE содержит 256 КБ собственной «локальной памяти» (local memory, также называемой local store или LS), разделенной на четыре отдельных сегмента по 64 КБ каждый, а также DMA-контроллер, который предназначен для обмена данными между основной памятью (RAM) и локальной памятью SPE (LS), минуя PPE. Доступ к LS составляет 6 тактов, что больше, чем время обращения к кэшу первого уровня, но меньше, чем к кэшу второго уровня для большинства современных процессоров. SPE-ядра, также как и PPE, используют упорядоченную схему (in-order) исполнения инструкций.

Частота всех ядер в процессоре Cell составляет 3,2 ГГц, что дает производительность одного SPE в  $3,2 \times 4 \times 2 = 25,6$  гигафлопс (последняя двойка в произведении за счет двух конвейеров, позволяющих за один такт выполнять операции умножения и сложения над вещественными числами). Таким образом, пиковая производительность всего процессора Cell получается превышающей 200 гигафлопс.

Модель программирования для процессора Cell «изначально» многопоточная, поскольку на SPE могут выполняться только специализированные потоки. Данные, с которыми они работают, должны располагаться в LS, соответственно типичным подходом является их предвыборка. В целом Cell весьма эффективно справляется с «поточковой» обработкой, характерной для мультимедиа, для задач кодирования, сжатия и т. д.

Основное отличие процессора PowerXCell™ 8i от своего «предка» состоит в значительном улучшении работы с вещественными числами двойной точности, что позволило довести пиковую производительность на них до уровня в 100 гигафлопс. Кроме того, PowerXCell™ 8i производится по 65 нм технологии, в отличие от 90 нм, использующихся в Cell. Наконец, в PowerXCell™ 8i был кардинально (до 32 GB) увеличен объем поддерживаемой памяти.

В ноябре 2012 г. в списке Top500 две системы были построены на процессорах PowerXCell 8i. Обе системы собраны в 2009 году. Как и ранее, приведем усредненные данные из списка Top500 по этим трем системам. Отношение «показанная мощность/пиковая мощность» систем на основе процессоров PowerXCell 8i составляет в 31-м списке 76,9%, при этом «удельная мощность» в расчете на одно ядро равна 11,24 гигафлопс (т. е. порядка 110 гигафлопс на процессор).

## **2.5. Процессоры Sun UltraSPARC T1, Sun UltraSPARC T2, SPARC T3, SPARC T4**

Начиная разработку микроархитектуры UltraSPARC Architecture, компания Sun Microsystems подошла к процессу с позиций, существенно отличающихся от остальных производителей. В многоядерных процессорах Intel, AMD и IBM каждое ядро фактически является полноценным исполнительным устройством, ориентированным на выполнение кода общего назначения, а в процессорах семейства Cell SPE-ядра, напротив, в принципе не могут исполнять такой код и, по сути, являются сопроцессорами.

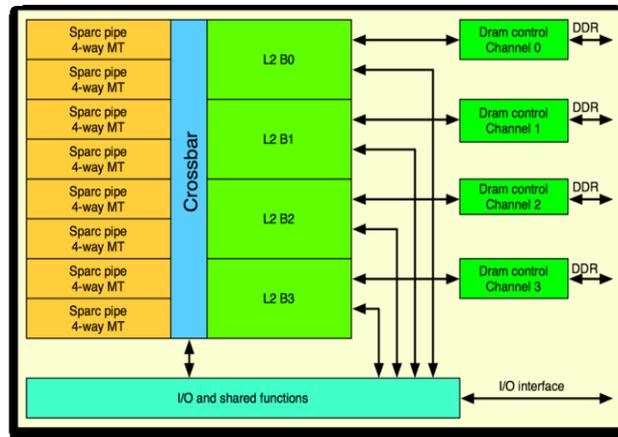


Рис. 7. Процессор Sun UltraSPARC T1

В основу процессоров UltraSPARC T1 (кодовое имя Niagara, см. рис. 1.9), выпущенных на рынок в 2005 г., UltraSPARC T2 (кодовое имя Niagara-2), выпущенных в 2007 г., SPARC T3, выпущенных в 2010 г., и SPARC T4, выпущенных в 2011 г., положена идея «многопоточности» для достижения высокой производительности не путем ускорения выполнения одного потока команд, а за счет обработки большого числа потоков в единицу времени. В результате процессоры UltraSPARC T1 способны выполнять 32 потока одновременно (на восьми «четырепоточных» ядрах), а процессоры UltraSPARC T2 – 64 потока (на восьми «восьмипоточных» ядрах). Эта многопоточность аппаратная (как, например, HyperThreading у компании Intel), т. е. операционная система воспринимает UltraSPARC T1 и UltraSPARC T2 как 32 и 64 процессора соответственно. Процессор SPARC T3 может содержать до 16 ядер по 8 потоков на каждом, а SPARC T4 – 8 ядер по 8 потоков на каждом.

В процессорах UltraSPARC T1 и T2 компания Sun реализовала технологию, названную ими CoolThreads, позволяющую значительно снизить энергопотребление – у процессоров UltraSPARC T1 не превышает 79 Вт (по 2,5 ватта на поток), а у процессоров UltraSPARC T2 – 123 Вт (всего 2 ватта на поток).

- Технические характеристики процессора Sun UltraSPARC T1:
- Тактовая частота: 1,0 или 1,2 ГГц.
- Число ядер: 8 (по 4 потока на каждое).
- Кэш инструкций первого уровня: 16 КБ на каждое ядро.
- Кэш данных первого уровня: 8 КБ на каждое ядро.
- Кэш второго уровня: 3 МБ (общий на все ядра).

- Интерфейс JBUS с пиковой пропускной способностью 3,1 GB/c, 128-битной шиной частотой от 150 до 200 МГц.
- 90-нм технологический процесс.
- Энергопотребление: 72 Вт, пиковое – 79 Вт.

Переключение между потоками в процессоре UltraSPARC T1 осуществляется по циклической схеме на каждом такте, т. е. в каждый конкретный момент времени активен только один из четырех потоков ядра. Однако в случае, если в потоке возникает простой (например, при кэш-промахе), ядро переключается на работу с другим потоком. Такая стратегия позволяет скрыть возникающие задержки доступа к памяти при наличии достаточного количества потоков исполнения. Ядра UltraSPARC T1 по функциональности аналогичны процессорам предыдущего поколения UltraSPARC III, но существенно упрощены архитектурно, например сокращены возможности прогноза ветвлений и спекулятивного выполнения команд, а число стадий конвейера уменьшено до шести (14 в UltraSPARC III).

Интересная особенность процессоров UltraSPARC T1 и T2 – наличие встроенного в ядро криптографического модуля (сопроцессора), реализующего на аппаратном уровне алгоритм RSA с 2048-разрядными ключами. Сопроцессор ядер в UltraSPARC T2 дополнительно поддерживает алгоритмы шифрования DES, 3DES, RC4, AES, SHA, MD5, CRC, а также алгоритм генерации случайных чисел.

Основной недостаток процессора UltraSPARC T1 – наличие в процессоре только одного блока вычислений над числами с плавающей точкой, доступного для всех потоков всех ядер. В процессоре UltraSPARC T2 эту проблему решили – у каждого ядра есть собственный модуль для выполнения операций с вещественными числами. Также в UltraSPARC T2 была поднята максимальная тактовая частота – до 1,4 ГГц. Плюс к этому – увеличен объем кэша второго уровня до 4 МБ, однако, в отличие от UltraSPARC T1, кэш не общий, а отдельный – по 512 КБ на каждое ядро. Кроме того, в процессор интегрированы два 10-Gbit контроллера Ethernet и контроллер шины PCI Express.

В процессоре SPARC T3 по сравнению с UltraSPARC T2 увеличен кэш второго уровня до 6 МБ, увеличена пропускная способность внешних интерфейсов, увеличено количество аппаратно поддерживаемых алгоритмов шифрования, увеличена максимальная тактовая частота до 1,67 ГГц. Производится процессор по технологии 40 нм. В одной системе может быть установлено до четырех процессоров SPARC T3.

В процессоре SPARC T4 реализовано внеочередное выполнение целочисленных команд, максимальная частота достигает 3 ГГц, а кэш 3-го уровня, размером 4 МБ, является общим для всех ядер.

### 3. Ускорители вычислений

Технологический мир сегодня пронизан конвергенцией – взаимным влиянием и даже взаимопроникновением технологий, стиранием границ между ними, возникновением многих интересных результатов на стыке областей в рамках междисциплинарных работ. Одно из проявлений этого явления – «игра» основных производителей аппаратных составляющих компьютеров на «чужих полях». Так, компании NVIDIA и ATI (последняя теперь – в составе компании AMD), накопив опыт и поняв, что пиковая мощность их продуктов уже стала сравнима с кластерами «средней руки», от выпуска графических ускорителей начали движение на рынок высокопроизводительных решений, представив соответствующие продукты (семейства NVIDIA® Tesla™ и ATI FireStream™). Напротив, компании, традиционно выпускавшие процессоры и серверные решения, взялись осваивать область мультимедиа: компания Intel разрабатывает многопоточные векторные графические устройства, компания IBM в составе альянса STI создала, как мы уже обсуждали, процессор Cell, изначально ориентированный именно на быструю обработку мультимедиа информации. Никуда не делись и типичные ускорители вычислений, способные существенно добавить мощности даже обычным «персоналкам» – на этом направлении работает, например, компания ClearSpeed Technology [108]. В конце 2012 года Intel представила свой сопроцессор с архитектурой Intel® Many Integrated Core, предназначенный для ускорения вычислений.

Как следствие перед обычными пользователями, желающими попробовать, «с чем едят» суперкомпьютерные технологии, возникает большой и не всегда просто осуществляемый выбор. Попробуем немного прокомментировать возможности трёх из представленных выше систем: ClearSpeed™ Advance™ e710, NVIDIA® Tesla™ K20X, Intel® Xeon Phi™ 5110P.

#### 3.1. Ускоритель ClearSpeed™ Advance™ e710

ClearSpeed™ Advance™ e710 (см. рис.) – это ускоритель операций над данными с плавающей запятой, представленных в формате с одинарной и двойной точностью. Ускоритель является сопроцессором, разработанным специально для серверов и рабочих станций, которые основаны на 32-х или 64-х битной x86 архитектуре, и построен на базе двухядерного процессора CSX700 содержащего по 96 обрабатывающих элемента на каждом ядре. Ускоритель e710 подключается к PCI-Express разъему на материнской плате. Среда разработки под e710 основана на языке C и включает SDK, а также набор инструментов для написания и отладки программ.



Рис. 8. Ускоритель ClearSpeed™ Advance™ e720<sup>9</sup>

Технические характеристики e710:

- 1 процессор CSX700, работающий на частоте 250 МГц.
- Каждый CSX700 содержит 192 обрабатывающих элемента (по 96 на каждое ядро).
- Пиковая производительность e710 – 150 гигафлопс.
- Память 2 GB DDR2-533 SDRAM.
- Пиковая пропускная способность внутренней памяти 192 GB/с.
- Пиковая пропускная способность доступа к внешней памяти 4 GB/с на ядро.
- Подключается к ПК через PCI-Express разъем материнской платы.
- Максимальная потребляемая мощность: 47 Вт.

Ускоритель e710 поддерживает операционные системы семейств Linux и Windows. Для работы с ускорителем необходимо установить специальный драйвер.

Главное (хост) приложение, использующее возможности e710, состоит из двух частей:

- программа, выполняющаяся на главном процессоре;
- программа, которая выполняется на ускорителе.

ClearSpeed™ Advance™ e710 обычно используется в качестве сопроцессора для ускорения внутренних циклов программы, достигающегося за счет использования стандартных математических библиотек, разработанных ClearSpeed. Если в приложении происходит вызов функции, поддерживаемой математическими библиотеками ClearSpeed, то библиотека сама анализирует возможность ее ускорения. В зависимости от выполненного анализа функция начинает выполняться либо на главном процессоре, либо пе-

---

<sup>9</sup> <http://www.clearspeed.com/products/advanceboard/e710.php>

рехватывается ускорителем. Этот механизм широко используется в ряде математических приложений, например в MATLAB.

Если в приложении происходит вызов функции, не поддерживаемой библиотеками ClearSpeed, то программист в случае необходимости может сам реализовать ее, используя имеющийся инструментарий.

### 3.2. GPU-ускорители NVIDIA® TESLA K20X

В 2007 г. компания NVIDIA представила продукты семейства Tesla™ для построения высокопроизводительных вычислительных систем. Последнее поколение ускорителей построено на базе архитектуры Kepler.

Ускоритель Tesla K20X содержит один процессор GK110, произведенный по технологии 28 нм. Ускоритель подключается к материнской плате по шине PCI Express 3.0.

Технические характеристики TESLA K20X:

- 2688 вычислительных ядер.
- Пиковая производительность на числах с плавающей запятой одинарной точности – 3,95 терафлопс.
- Пиковая производительность на числах с плавающей запятой двойной точности – 1,31 терафлопс.
- Память 6 GB GDDR5 384-bit.
- Пиковая пропускная способность памяти 250 GB/c.

Tesla K20X поддерживает операционные системы семейств Windows и Linux. В комплект поставки входят следующие компоненты: CUDA Driver, CUDA Toolkit, CUDA SDK.

CUDA (Compute Unified Device Architecture) – программно аппаратное решение, позволяющее использовать видеопроцессоры для вычислений общего назначения. Возможность программирования видеопроцессоров компании NVIDIA на CUDA существует, начиная с семейства видеокарт 8xxx. Разработка программ для выполнения на Tesla K20X происходит таким же способом.

С точки зрения программиста, Tesla K20X представляет собой набор независимых мультипроцессоров. Каждый мультипроцессор состоит из нескольких независимых скалярных процессоров, двух модулей для вычисления математических функций, конвейера, а также общей памяти.

CUDA позволяет создавать специальные функции (ядра, kernels), которые выполняются параллельно различными блоками и потоками, в отличие от обычных C-функций. При запуске ядра блоки распределяются по доступ-

ным мультипроцессорам. Мультипроцессор занимается распределением, параллельным выполнением потоков внутри блока и их синхронизацией. Каждый поток независимо исполняется на одном скалярном процессоре с собственным стеком инструкций и памятью.

CUDA предоставляет программисту доступ к нескольким уровням памяти:

- каждый поток обладает локальной памятью;
- все потоки внутри блока имеют доступ к быстрой общей памяти блока, время жизни которой совпадает со временем жизни блока; память блока разбита на страницы, доступ к данным на разных страницах осуществляется параллельно;
- все потоки во всех блоках имеют доступ к общей памяти устройства.

Всем потокам также доступны два вида общей памяти для чтения: константная и текстурная, они кэшируются. Так же, как и в общей памяти устройства, данные сохраняются на протяжении работы приложения.

В ноябре 2012 г. в списке Top500 две системы были построены с использованием ускорителей Tesla K20X, среди них лидер этого списка – Titan.

### 3.3. Сопроцессор Intel® Xeon Phi™ 5110P

В конце 2012 года Intel представила первый процессор с архитектурой Intel MIC (Intel® Many Integrated Core Architecture). Основой архитектуры MIC является использование большого количества вычислительных ядер архитектуры x86 в одном процессоре. В результате для разработки параллельных программ могут быть использованы стандартные технологии<sup>10</sup>: pthreads, OpenMP, Intel TBB, Intel Cilk Plus, MPI. Кроме того, под архитектуру x86 разработано огромное количество инструментов и библиотек. Это является огромным преимуществом перед другими ускорителями вычислений, для которых используются специальные (обычно более сложные) технологии разработки параллельных программ: CUDA, OpenCL.

Intel® Xeon Phi™ 5110P подключается к разъему PCIe x16 на материнской плате. В двухпроцессорной системе можно установить до 8 карт Intel® Xeon Phi™ 5110P.

---

<sup>10</sup> Заметим, что использование указанных технологий невозможно без использования компилятора Intel C++ Compiler или Intel Fortran Compiler.

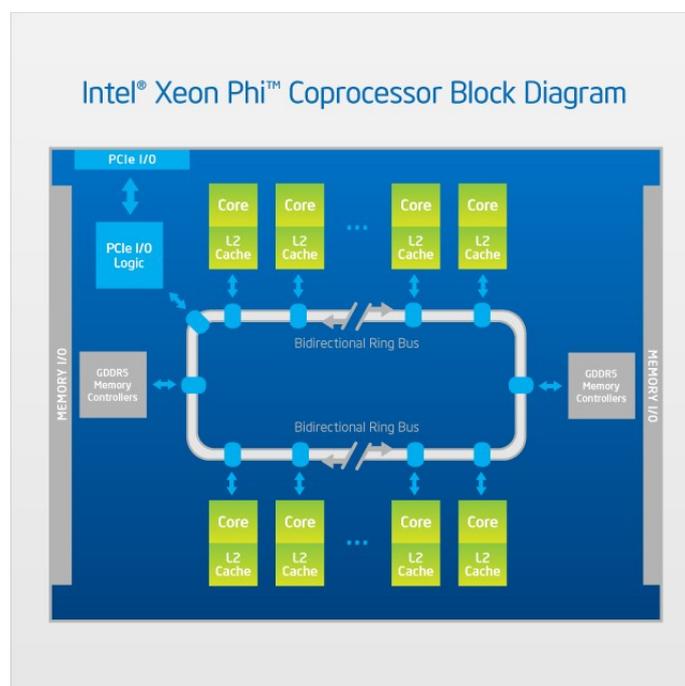


Рис. 9. Архитектура Intel® Xeon Phi™ 5110P <sup>11</sup>

Технические характеристики Intel Xeon Phi 5110P:

- 60 вычислительных ядер.
- 240 потоков (по 4 потока на ядро).
- Частота ядер – 1,053 ГГц.
- Пиковая производительность на числах с плавающей запятой двойной точности – 1.0108 терафлопс.
- Кэш второго уровня: 30 МБ (по 512 КБ на каждое ядро).
- Память 8 GB GDDR5.
- Пиковая пропускная способность памяти 320 GB/с.
- Максимальная потребляемая мощность: 225 Вт.

Каждое ядро может обрабатывать одновременно 4 потока и поддерживает векторные инструкции с размером операндов 512 бит (инструкции MMX, SSE, AVX не поддерживаются). Все ядра подключены к двунаправленной кольцевой шине.

<sup>11</sup> <http://www.intel.com/content/www/us/en/processors/xeon/xeon-phi-coprocessor-block-diagram.html>

Intel® Xeon Phi™ 5110P поддерживает операционные системы семейства Linux и прямую ip-адресацию, что упрощает интеграцию сопроцессора в кластер.