

МНОГОЯДЕРНОСТЬ, КАК СПОСОБ УВЕЛИЧЕНИЯ ПРОИЗВОДИТЕЛЬНОСТИ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ

Э.П. Ланина

Иркутский государственный технический университет
664074, г. Иркутск, ул. Лермонтова, д. 83

Аннотация. Существует два подхода к увеличению производительности процессора. Первый - увеличение тактовой частоты процессора, второй – увеличение количества инструкций программного кода, выполняемых за один такт процессора. Увеличение тактовой частоты не может быть бесконечным и определяется технологией изготовления процессора. При этом рост производительности не является прямо пропорциональным росту тактовой частоты, то есть наблюдается тенденция насыщаемости, когда дальнейшее увеличение тактовой частоты становится нерентабельным. Разработка более совершенных архитектур процессоров, содержащих большее число функциональных исполнительных устройств, с целью повышения количества команд, одновременно исполняемых за один такт, — традиционный альтернативный росту тактовой частоты путь повышения производительности. Но такие разработки очень сложны и дороги. Сложность разработки возрастает с ростом сложности логики экспоненциально.

Можно сказать, что идея построения многоядерных микропроцессоров является развитием идеи кластеров, но в данном случае дублируется целиком процессорное ядро. Другим предшественником многоядерного подхода можно считать технологию Intel - HyperThreading, где также есть небольшое дублирование аппаратуры и использование двух потоков инструкций, использующих общее ядро.

Многоядерный процессор имеет два или больше "исполнительных ядер". Операционная система рассматривает каждое из исполнительных ядер, как дискретный процессор со всеми необходимыми вычислительными ресурсами. Поэтому многоядерная архитектура процессора, при поддержке соответствующего

программного обеспечения, осуществляет полностью параллельное выполнение нескольких программных потоков.

К 2006 году все ведущие разработчики микропроцессоров создали двоядерные процессоры. Переход к многоядерным процессорам становится основным направлением повышения производительности вычислительных систем. В связи с этим, знание основ функционирования вычислительных систем на многоядерных процессорах является актуальным.

Annotation. There are two ways for increasing processor performance. The first one is to increase CPU Frequency, the second is to increase the amount of Instruction Per Clock. The increment of CPU frequency can not be infinite and it depends of processor production technology. At the same time the throughput growth is not straight proportional to increment of CPU frequency that is tendency of saturating is observed when further CPU increment became uneconomic. Traditionally the alternative way for CPU increment is the development of more efficient processor architectures that contain more functional executive units to increase the quantity of instructions that are executed simultaneously in one circle. But such developments are very difficult and expensive. The development complexity grows exponentially during logic complexity increase.

Could say the multicore processor idea is the development of cluster idea but in this case the whole core is duplicated. The Intel HyperThreading technology might be considered as another predecessor of multicore approach. There is also little equipment duplication in HyperThreading technology and it exploits two threads of instructions that utilize one core.

The multicore processor has two or more “executive cores”. Operating system distinguishes every executive core as a discrete processor that has all necessary computing resource. Therefore the multicore processor architecture implements fully parallel execution of several program streams. All principal processor developers created double core processors at 2006. Migration to multicore processors technology becomes the main trend for increment of systems performance. In this connection the basic knowledge of computer systems with multicore processors operation nowadays is actual.

1. Введение

Закон Мура гласит, что количество транзисторов, размещенных на полупроводниковой микросхеме, удваивается каждые два года, что приводит, с одной стороны, к повышению производительности, а с другой стороны, к снижению стоимости производства микросхем. Несмотря на важность и действенность этого закона в течение долгих лет, оценивая перспективы дальнейшего развития, время от времени предсказывали его неминуемое фиаско.

В качестве препятствий на пути дальнейшего развития называются такие факторы, как ограничения из-за физических размеров, стремительный рост энергопотребления и непомерно высокие затраты на производство.

На протяжении многих лет для повышения производительности процессора использовался единственный путь - повышение его тактовой частотой. За эти годы укоренилось мнение, что именно тактовая частота процессора является основным показателем его производительности. Нарастивание тактовых частот на современном этапе не простая задача. Конец гонке тактовых частот микропроцессоров был положен благодаря нерешенной проблеме токов утечки и неприемлемому росту тепловыделения микросхем.

Производительность процессора (Performance) — это отношение общего числа выполненных инструкций программного кода ко времени их выполнения или количество инструкций, выполняемых за секунду (Instructions rate):

$$\text{производительность} = \frac{\text{количество инструкций}}{\text{время выполнения}}$$

Так как основной характеристикой процессора стала его тактовая частота, то введем частоту в формулу производительность процессора. Помножим числитель и знаменатель на количество тактов, за которое выполнены инструкции:

$$\text{производительность} = \frac{\text{количество инструкций}}{\text{количество тактов}} = \frac{\text{количество тактов}}{\text{время выполнения}}$$

Первая часть полученного произведения - количество инструкций, выполняемых за один такт (Instruction Per Clock, IPC), Вторая часть произведения - количество тактов процессора в единицу времени (тактовая частота процессора, F или Frequency). Таким образом, производительность процессора зависит не только от его тактовой частоты, но и от количества инструкций, выполняемых за такт (IPC):

$$\text{производительность} = (\text{IPC}) (\text{F})$$

Полученная формула определяет два разных подхода к увеличению производительности процессора. Первый - увеличение тактовой частоты процессора, а второй – увеличение количества инструкций программного кода, выполняемых за один такт процессора. Увеличение тактовой частоты не может быть бесконечным и определяется технологией изготовления процессора. При этом рост производительности не является прямо пропорциональным росту тактовой частоты, то есть наблюдается тенденция насыщаемости, когда дальнейшее увеличение тактовой частоты становится нерентабельным. Количество инструкций, выполняемых за время одного такта, зависит от микроархитектуры процессора: от количества исполнительных блоков, от длины конвейера и эффективности его заполнения, от блока предвыборки, от оптимизации программного кода к данной микроархитектуре процессора. Поэтому сравнение производительности процессоров на основании их тактовой частоты возможно только в пределах одной и той же архитектуры (при одинаковом значении количества выполняемых операций в секунду - IPC процессоров).

Сравнение производительности процессоров с различной архитектурой на основе тактовой частоты неправомерно. К примеру, основываясь на тактовой частоте, некорректно сравнивать производительности процессоров с разным размером кэш памяти уровня L2, или производительности процессоров, поддерживающих и не поддерживающих технологию Hyper-Threading.

Отход от последовательного исполнения команд и использование нескольких исполняющих блоков в одном процессоре позволяют одновременно обрабатывать несколько процессорных микрокоманд, то есть организовывать параллелизм на уровне инструкций (Instruction Level Parallelism - ILP), что, разумеется, увеличивает общую производительность.

Еще один подход к решению данной проблемы был реализован в VLIW/EPIC - архитектуре IA-64 (очень длинных команд), где часть проблем переложена с аппаратуры на компилятор. И все же разработчики признают, что для достижения высокой производительности архитектура важнее.

При большом числе функциональных блоков микросхемы и большом ее размере возникает проблема, связанная со скоростью распространения сигналов - за один такт сигналы не успевают достигнуть необходимые блоки. В качестве возможного выхода в микропроцессорах Alpha были введены так называемые "кластеры", где устройства частично дублировались, но зато внутри кластеров расстояния были меньше. Можно сказать, что идея построения многоядерных микропроцессоров является развитием идеи кластеров, но в данном случае дублируется целиком процессорное ядро.

Другим предшественником многоядерного подхода можно считать технологию Intel - HyperThreading, где также есть небольшое дублирование аппаратуры и использование двух потоков инструкций, использующих общее ядро.

Многоядерный процессор имеет два или больше "исполнительных ядер". Ядром процессора можно назвать его систему исполнительных устройств (набор арифметико-логических устройств), предназначенных для обработки данных. Операционная система рассматривает каждое из исполнительных ядер, как дискретный процессор со всеми необходимыми вычислительными ресурсами. Поэтому многоядерная архитектура процессора, при поддержке соответствующего программного обеспечения, осуществляет полностью параллельное выполнение нескольких программных потоков.

К 2006 году все ведущие разработчики микропроцессоров создали двуядерные процессоры. Первыми появились двуядерные RISC-процессоры Sun Microsystems (UltraSPARC IV), IBM (Power4, Power5) и HP (PA-8800 и PA-8900).

О выпуске двуядерных процессоров с архитектурой x86 фирмы AMD и Intel объявили почти одновременно.

Архитектура процессоров достигла достаточно высокой сложности, поэтому переход к многоядерным процессорам становится основным направлением повышения производительности вычислительных систем.

2. Многоядерные процессоры и новые технологии многоядерной обработки данных

2.1. Технологии многопоточности

Основные направления развития архитектуры современных микропроцессоров определяются стремлением к увеличению их производительности. Последнюю можно повышать, например, наращивая тактовую частоту и (или) увеличивая число команд, выполняемых за один такт. Одно из решений данной проблемы связано с реализацией концепции "параллелизма на уровне тредов (поток)" - TLP (Thread Level Parallelism). Если программные коды не в состоянии загрузить работой все или даже большинство функциональных устройств, то можно разрешить процессору выполнять более чем одну задачу (тред, или поток), чтобы дополнительные потоки загрузили простаивающие устройства. Здесь нетрудно усмотреть аналогию с многозадачной операционной системой: чтобы процессор не простаивал, когда задача оказывается в состоянии ожидания (например, завершения ввода-вывода), операционная система переключает процессор на выполнение другой задачи. Более того, некоторые механизмы диспетчеризации в операционной системе (например, квантование) имеют аналоги в многопоточковой архитектуре (МТА - MultiThreading Architecture). Очевидно, архитектура, поддерживающая параллелизм на уровне потоков (TLP), должна гарантировать, что треды не будут использовать одновременно одни и те же ресурсы, для чего требуются дополнительные аппаратные средства (например, дублирование регистровых файлов). Однако оказалось, что можно реализовать МТА на базе современных суперскалярных процессоров, и это требует лишь относительно небольших аппаратных доработок, что резко повышает привлекательность МТА в глазах проектировщиков.

При использовании базового типа параллелизма на уровне потоков (TLP) в микропроцессоре необходимо иметь не менее двух аппаратных расширений для потоков. Это регистры общего назначения, счетчик команд, слово состояния процесса и т. п. В любой момент времени работает только один поток (тред). Он выполняется до возникновения определенной ситуации (например, выполнения команды загрузки

регистра при отсутствии данных в кэш-памяти). В этом случае процессор переключается на выполнение другого потока. Поскольку при непопадании в кэш-память операции с памятью могут потребовать до сотни тактов процессора, его простои по причине ожидания данных могли бы быть весьма значительными. Современные процессоры, имеющие возможности спекулятивного внеочередного выполнения команд, в подобной ситуации могут продолжить выполнение других команд, но на практике число независимых команд быстро исчерпывается и процессор останавливается.

Архитектура с одновременным выполнением тредов - SMT (Simultaneous Multi-Threading) допускает одновременное выполнение нескольких потоков. В этом случае на каждом новом такте на выполнение в какое-либо исполнительное устройство может направляться команда любого потока. По сравнению с суперскалярными процессорами, поддерживающими внеочередное спекулятивное выполнение команд и использующими механизм переименования регистров, для SMT необходимы, в частности, следующие аппаратные средства:

- несколько счетчиков команд (по одному на поток) с возможностью выбора любого из них на каждом такте;
- средства, ассоциирующие команды с потоком, которому они принадлежат (необходимы, в частности, для работы механизмов предсказания переходов и переименования регистров);
- несколько стеков адресов возврата (по одному на поток) для предсказания адресов возврата из подпрограмм;
- специальная дополнительная память в процессоре (в расчете на каждый поток) для процедуры удаления из буфера выполненных вне очереди команд.

Одна из основных особенностей SMT у многих современных процессоров - переименование регистров, когда логические (архитектурные) регистры отображаются в физические, с которыми и ведется реальная работа. Техника переименования регистров может, очевидно, применяться для того, чтобы избежать прямого дублирования файлов регистров, как аппаратной принадлежности потока.

2.2. Технология Hyper-Threading

Анонсированная в 2002 году компанией Intel технология Hyper-Threading – пример многопоточной обработки команд. Данная технология является чем-то средним между многопоточной обработкой, реализованной в мультипроцессорных системах, и параллелизмом на уровне инструкций, реализованном в однопроцессорных системах. Фактически технология Hyper-Threading позволяет организовать два логических процессора в одном физическом. Таким образом, с точки зрения операционной системы и запущенного приложения в системе существует два процессора, что даёт возможность распределять загрузку задач между ними.

Посредством реализованного в технологии Hyper-Threading принципа параллельности можно обрабатывать инструкции в параллельном (а не в последовательном) режиме, то есть для обработки все инструкции разделяются на два параллельных потока. Это позволяет одновременно обрабатывать два различных приложения или два различных потока одного приложения и тем самым увеличить IPC процессора (количество инструкций, выполняемых процессором в секунду), что сказывается на росте его производительности.

В конструктивном плане процессор с поддержкой технологии Hyper-Threading состоит из двух логических процессоров, каждый из которых имеет свои регистры и контроллер прерываний (Architecture State, AS). А значит, две параллельно исполняемые задачи работают со своими собственными независимыми регистрами и прерываниями, но при этом используют одни и те же ресурсы процессора для выполнения своих задач. После активизации каждый из логических процессоров может самостоятельно и независимо от другого процессора выполнять свою задачу, обрабатывать прерывания либо блокироваться. Таким образом, от реальной двухпроцессорной конфигурации новая технология отличается только тем, что оба логических процессора используют одни и те же исполняющие ресурсы, одну и ту же разделяемую между двумя потоками кэш-память и одну и ту же системную шину. Использование двух логических процессоров позволяет усилить процесс параллелизма на уровне потока, реализованный в современных операционных системах и высокоэффективных приложениях. Команды от исполняемых параллельно потоков

одновременно посылаются для обработки ядру процессора. Используя технологию out-of-order (исполнение командных инструкций не в порядке их поступления), ядро процессора способно параллельно обрабатывать оба потока за счёт использования нескольких исполнительных модулей.

Идея технологии Hyper-Threading тесно связана с микроархитектурой NetBurst процессора Pentium 4 и является в каком-то смысле её логическим продолжением.

Микроархитектура Intel NetBurst позволяет получить максимальный выигрыш производительности при выполнении одиночного потока инструкций, то есть при выполнении одной задачи. Однако, даже в случае специальной оптимизации программы, не все исполнительные модули процессора оказываются задействованными на протяжении каждого тактового цикла. В среднем, при выполнении кода, типичного для набора команд IA-32, реально используется только 35% исполнительных ресурсов процессора, а 65% исполнительных ресурсов процессора простаивают, что означает неэффективное использование возможностей процессора. Было бы логично реализовать работу процессора таким образом, чтобы в каждом тактовом цикле максимально использовать его возможности. Именно эту идею и реализует технология Hyper-Threading, подключая незадействованные ресурсы процессора к выполнению параллельной задачи.

В современных приложениях в любой момент времени, как правило, выполняется не одна, а несколько задач или несколько потоков (тредов - threads) одной задачи, называемых также нитями. В качестве примера рассмотрим работу двух потоков. При одновременном выполнении обоих потоков процессор будет постоянно переключаться между потоками, за один такт процессора выполняются только инструкции какого-либо одного из потоков. На каждом такте процессора используются далеко не все исполнительные блоки процессора, поэтому имеется возможность частично совместить выполнение инструкций отдельных потоков на каждом такте процессора. Например, выполнение двух арифметических операций с целыми числами первого потока можно совместить с загрузкой данных из памяти второго потока и выполнить все три операции за один такт процессора. Аналогично на втором такте процессора можно совместить операцию сохранения результатов первого потока с двумя операциями второго потока и т.д. Собственно, в таком параллельном

выполнении двух потоков и заключается основная идея технологии Hyper-Threading. Hyper-Threading - это виртуальная многопроцессорность, так как процессор на самом деле один, а операционная система видит два процессора. Классическому "одноядерному" процессору добавили еще один логический блок управления AS (Architectural State) технологии IA-32 (рис.1).



Рис.1. Блоки управления AS (Architectural State) технологии IA-32.

Architectural State отслеживает состояние регистров (общего назначения, управляющих, прерываний - APIC, служебных), используя единственное физическое ядро (блоки предсказания ветвлений, ALU, FPU, SIMD-блоки и пр.) AS1 представляет из себя один логический процессор (LP1), AS2 — второй логический процессор (LP2). У каждого логического процессора есть свой собственный контроллер прерываний (APIC — Advanced Programmable Interrupt Controller) и набор регистров. Для корректного использования общих регистров двумя логическими процессорами существует специальная таблица — RAT (Register Alias Table), согласно которой можно установить соответствие между регистрами общего назначения общего процессорного ядра. Таблица использования регистров RAT у каждого логического процессора своя. В результате получается схема, при которой на одном и том же ядре могут свободно выполняться два независимых потока программного кода.

У процессора с технологией Hyper-Threading предусмотрены два основных режима работы: Single-Task (ST) и Multi-Task (MT). В режиме ST активным является только один логический процессор, который безраздельно пользуется доступными

ресурсами, другой логический процессор остановлен командой HALT. При появлении второго программного потока бездействовавший логический процессор активируется (посредством прерывания), и физический процессор переводится в режим работы с двумя потоками (рис.2).

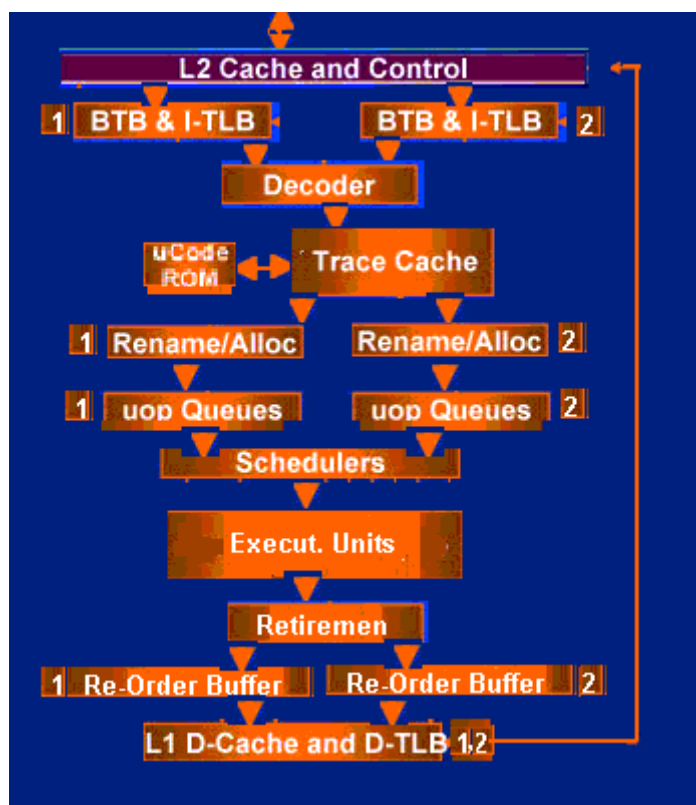


Рис.2. Разделение ресурсов двумя логическими процессорами.

При работе двух потоков поддерживаются две очереди инструкций (Next Instruction Pointers). Большая часть инструкций берется из кэш трассировки Trace Cache (TC), где они хранятся в декодированном виде, и доступ к кэш трассировки два активных логических процессора получают поочередно, через такт. В то же время, когда активен только один логический процессор, он получает монопольный доступ к кэш трассировки без чередования по тактам. Аналогичным же образом происходит и доступ к памяти микропрограмм (Microcode ROM). Блоки трансляции инструкций (ITLB - Instruction Translation Look-aside Buffer) действуют при отсутствии необходимых инструкций в кэш команд и доставляют команды, каждый для своего

потока. Блок декодирования инструкций (IA-32 Instruction Decode) является разделяемым и в случае, когда требуется декодирование инструкций для обоих потоков, обслуживает их поочередно (через такт). Блоки очереди декодированных команд (Uop Queue) и блок распределения команд по исполнительным устройствам (Allocator) разделяются надвое, отводя по половине элементов для каждого логического процессора. Пять планировщиков (Schedulers) обрабатывают очереди декодированных команд, независимо от принадлежности их к какому либо потоку, и направляют команды на выполнение нужным исполнительным устройствам (Execution Units) в зависимости от готовности к выполнению команд и доступности исполнительных устройств. Кэш память всех уровней является полностью разделяемой между двумя логическими процессорами, но для обеспечения целостности данных записи в буфер данных (DTLB - Data Translation Look-aside Buffer) снабжаются дескрипторами в виде идентификатора (ID) каждого логического процессора.

Начальная часть конвейера Pentium 4 отвечает за подачу микроопераций (декодированных x86 инструкций) на исполнительную часть конвейера. Именно тут, в основном, и находятся дублирующиеся блоки для каждого из двух логических процессоров. Кэш трассировки (Trace Cache) содержит уже декодированные инструкции. Большинство команд при реальной работе процессора декодировано заранее и находится в кэш трассировки. Trace Cache не дублируется для каждого из логических процессоров, а разделяется между ними. Тем не менее, у каждого логического процессора существует свой собственный блок трассировки, ссылающийся на следующую инструкцию для выполнения. Инструкции из Trace Cache выбираются по очереди и становятся в так называемую очередь выборки, также индивидуальную для обоих логических процессоров.

При отсутствии очередной инструкции в Trace Cache, являющемся в Pentium 4 по своей иерархии кэш памятью первого уровня для инструкций, процессор должен декодировать очередную x86 инструкцию из L2 - кэш второго уровня. Выборка инструкции осуществляется при участии блока трансляции адреса команд (Instruction Translation Lookaside Buffer- ITLB), переводящем адрес, хранящийся в блоке трассировки, в физический адрес. ITLB является также индивидуальным для каждого процессора, а L2-кэш разделяется между логическими CPU. Декодер x86 инструкций в

процессорах с технологией Hyper-Threading общий, поскольку его загрузка невелика – большинство декодированных инструкций уже хранится в Trace Cache. Если же оба логических процессора обращаются к декодеру одновременно, он чередует свою работу между логическими процессорами, но только после выполнения полного цикла декодирования для одного из логических CPU. Декодированные инструкции записываются в Trace Cache.

К исполнительному устройству декодированные последовательности инструкций приходят в двух очередях – своя очередь для каждого из логических процессоров. В первую очередь инструкции из двух входящих очередей проходят через блоки выделения ресурсов(Allocator) и переименования регистров(Register Rename). Здесь происходит выделение необходимых ресурсов для исполнения команд. Машинные регистры и буферы делятся поровну между логическими процессорами, однако, при отказе одного из логических процессоров от использования тех или иных ресурсов, они могут быть всецело выделены только одному логическому процессору. После прохождения этой стадии команды попадают в две отсортированные очереди – для операций с памятью и для остальных операций, которые также разделены пополам – для каждого из логических процессоров.

Затем, рассортированные микрооперации приходят на стадию определения очередности выполнения (Scheduling), где выполняется сортировка порядка следования инструкций при поступлении на исполнительные устройства. Операции на блоки-планировщики приходят по мере поступления. При необходимости, планировщики переключаются с очереди одного логического процессора на очереди другого. На этом этапе, кстати, происходит окончательное смешение микроопераций, приходящих с логических процессоров, для возможности их одновременного исполнения. Поскольку машинные регистры физического процессора к этому моменту оказываются жестко привязанными к регистрам обоих логических процессоров, исполнение инструкций действительно становится возможным без разбора принадлежности команд.

После этапа выполнения, на котором процессор не различает логические процессоры, следует блок восстановления (Retirement), где восстанавливается изначальный порядок инструкций и их принадлежность к каждому из логических

процессоров. При этом буфер восстановления порядка (Re-Order Buffer) делится пополам между логическими процессорами.

Хотя кэш память первого и второго уровня является разделяемой между логическими процессорами, буфер трансляции данных (Data Translation Lookaside Buffer - DTLB), сопоставляющий адреса данных и их физические адреса, хоть и делится между процессорами, но записи в нем дополнены идентификатором процессора, которому принадлежит каждая из строк буфера.

Таким образом, технология Hyper-Threading действительно позволяет загрузить исполнительные устройства процессора значительно сильнее за счет одновременного выполнения двух потоков. Однако, следует понимать, что эффект от такого приема не всегда может быть положительным. Во-первых, если выполняемые потоки похожи по типу выполняемых инструкций, выигрыша может не быть вовсе, поскольку один из потоков будет полностью занимать все ресурсы, необходимые и другому потоку. Простой же остальных исполнительных устройств процессора от этого не исчезнут. Во-вторых, возможна и куда более катастрофичная ситуация. Один из потоков может попросту занять ресурсы, необходимые другому потоку, и ожидать при этом, например, поступления данных. Операционная система же при этом, пребывая в уверенности, что имеется два CPU, не будет предпринимать никаких действий, на самом же деле функционирование процессора будет попросту парализовано. Именно поэтому Intel обратилась с призывом к разработчикам программного обеспечения оптимизировать свои программы для Hyper-Threading. Один из основных принципов такой оптимизации – применение новой инструкции PAUSE, не допускающей пустых циклов ожидания. В ноябре 2002 года, компания Intel официально анонсировала свой очередной процессор в семействе Pentium 4, Intel Pentium 4 3.06 ГГц. Этот процессор являлся первым CPU в семействе, поддерживающем технологию Hyper-Threading.

2.3. Многоядерность

Для реализации процесса параллельного выполнения задач более эффективно интегрировать два ядра или более в одном микропроцессоре. Такая многоядерная конфигурация на одном кристалле обеспечивает более высокую скорость обмена

между ядрами, чем использование внешних шин, коммутаторов и т.п. в многопроцессорных системах.

Многоядерная архитектура предоставляет два или более полнофункциональных набора ресурсов для повышения производительности процессора. Многоядерность и 65нм техпроцесс позволили добиться значительной экономии энергопотребления и повышения производительности на 1 Вт потребляемой мощности.

Совместно с многоядерными процессорами, в архитектуру новых платформ внедряются такие новые технологии, как независимость соответствующих программных компонент (Intel Virtualization Technology - VM), ускорение механизма обмена данными (Intel I/O Acceleration Technology - I/O AT), удаленная управляемость (Intel Active Management Technology - iAMT).

Комплекс новых технологий направлен на повышение эффективности вычислительной платформы в целом.

2.4. Технология виртуализации Hypervisor

Virtualization Technology Hypervisor является, по сути, сочетанием новых аппаратных возможностей многоядерных процессоров будущих поколений, реализующих виртуализацию (независимость соответствующих программных компонент).

Процессоры, поддерживающие технологию виртуализации Hypervisor, понимают дополнительные инструкции, позволяющие параллельную работу нескольких ОС, и позволяют между ними переключаться.

Существуют программные решения, которые позволяют запускать параллельно несколько операционных систем (виртуальных машин), однако аппаратная поддержка позволяет ускорить этот процесс и сделать его более надёжным. Таким образом, достигается защищенность и независимость выполнения разных процессов на одной системе, когда даже критический сбой одного процесса никак не может повлиять на выполнение других процессов.

Реализация этой технологии требует поддержки в BIOS, разработки специальной системы управления (монитора виртуальных машин VMM) и поддержки со стороны прикладных программ.

2.5. Технология удаленного управления компьютерами сети

Active Management Technology (AMT) - технология удаленного управления компьютерами сети. Эта технология позволяет администраторам сети производить удаленное администрирование и обслуживание сетевых компьютеров на качественно новом уровне. В перечне доступных возможностей представлены: удаленная диагностика и установление неисправностей оборудования, удаленная установка и обновление программного обеспечения, защита от вирусов.

2.6. Технология ускорения процесса ввода-вывода данных - Acceleration Technology

Реализация этой технологии основана на аппаратных возможностях процессора нового поколения с оптимизированным стеком протоколов, на возможностях новых типов чипсетов и сетевых адаптеров с оптимизированным доступом к данным, что позволит добиться оптимизации скорости передачи данных.

Технология ускорения ввода-вывода I/OAT (Intel I/O Acceleration Technology) появилась в то время, когда потребности приложений (таких, как электронная коммерция, обмен сообщениями, приложения для кластеров устройств хранения данных и серверов) начали обгонять способность серверов к их возможности быстрого и надежного обмена сетевыми данными с приложениями. В то время как производительность серверных процессоров и пропускная способность сетей значительно выросли, основной метод обмена данными остался прежним. Процессор сервера ведет обработку данных, осуществляет доступ к памяти и реализует протоколы обмена для каждого пакета данных. В результате работа серверных приложений замедляется.

Для решения этой проблемы в технологии ускорения ввода-вывода корпорацией

Intel применяется общеплатформенный подход. Задача управления данными распределяется между всеми компонентами платформы - процессором, набором микросхем, сетевым контроллером и программным обеспечением. Общеплатформенный подход позволяет снизить нагрузку на процессор и ускорить обмен данными. Загруженность процессора снижается благодаря тому, что набор микросхем и сетевой контроллер получают возможность считывать данные из памяти и записывать их в память.

Intel также оптимизировала протокол TCP/IP - открытый "свод правил", который позволяет компьютерам всех типов обмениваться данными, общаясь на одном языке. В результате загруженность процессоров в серверах архитектуры Intel снизилась наполовину, а вычислительные ресурсы высвободились для решения других задач. В среднем такой подход позволяет ускорить обмен данными между платформой и приложениями на 30% и освобождает процессор для выполнения другой вычислительной работы.

Кроме того, подход, применяемый в технологии I/OAT, позволяет отказаться от искусственных надстроек, применяемых в существующих технологиях, - таких, как механизмы разгрузки TCP (TCP offload engine, TOE). TOE - это специализированные дорогостоящие микросхемы, предназначенные для разгрузки процессора при обработке протокола TCP/IP, но они не решают двух основных проблем, связанных с процессором: снижения системных издержек и организации доступа к памяти. В результате TOE эффективны только для таких приложений, где велика информационная составляющая пакетов данных - например, для высокопроизводительных систем управления базами данных или для хранилищ данных.

Корпорация Microsoft обещает обеспечить встроенную поддержку технологии I/OAT в будущих версиях ОС Windows Server. В этих версиях также будет использоваться технология, которая позволяет сбалансировать трафик TCP/IP при использовании многоядерных процессоров.

Выводы:

Внедрение новых технологий:

- Улучшает возможности работы в многозадачных средах с одновременным выполнением нескольких активных и фоновых приложений.
- Обеспечивает меньшее время загрузки и более высокую производительность для потоковых приложений
- Уменьшает время отклика для совместного запуска приложений, повышает возможности многозадачных приложений.
- Позволяет выполнять параллельно нескольких операций с высокой пропускной способностью, при этом возможно увеличить количество пользователей, работающих одновременно на одном ПК.
- Позволяет усовершенствовать возможности для совместной работы.
- Позволяет повысить эффективность и снизить энергопотребление при одновременном запуске множества приложений

Фактически пластину с выращенными ядрами процессоров можно разрезать и по одному ядру, и по парам. Режим работы уже готового процессора определяется коммутацией в упаковке. Так что, выпуская всего одно ядро, можно делать весь спектр процессоров — от одно до многоядерных. Такой подход может и не стать основным, поскольку имеет недостатки. Прежде всего, это неоптимальное использование кэш-памяти. Независимые процессоры с большой вероятностью кэшируют одни и те же данные, и при их модификации необходим механизм взаимного уведомления процессоров (обеспечить когерентность кэш памяти).

Очевидно, что кэш-память должна быть общей для всех ядер. Однако собрать такой процессор из нескольких независимых заготовок невозможно. Тут необходима разработка абсолютно нового ядра — а это сложно и дорого.

3. Многоядерные процессоры корпорации Intel

Корпорация Intel (www.intel.com) давно работает над концепцией параллелизма и аппаратными средствами реализации многопоточности. Уже к 1994 г. в процессоре Intel Pentium был реализован параллелизм на уровне команд - архитектурная особенность, при которой команды одного потока кода извлекаются, выполняются параллельно, а затем объединяются в прежнем порядке. В 1994 г. корпорация реализовала также двухпроцессорную обработку (два полноценных процессора, вставлялись в два разъема системной платы), создав аппаратную многопоточную среду для серверов и рабочих станций. В 1995 г. был представлен процессор Pentium Pro, поддерживавший эффективное объединение уже четырех процессоров на одной системной плате, что обеспечило более высокую скорость обработки данных в многопоточных приложениях, ориентированных на серверные платформы и рабочие станции.

Эти усилия стимулировали разработку однопроцессорных технологий, обеспечивающих более высокую степень параллелизма на уровне потоков, для массовых платформ. Корпорация реализовала технологию НТ (Hyper-Threading) для процессоров Pentium 4 и Xeon, как инновационный способ достижения более высокой степени параллелизма на уровне потоков в процессорах для массовых систем. В корпорации поняли, что технология НТ естественным образом ведет к разработке многоядерных процессоров, отличающихся более высокими степенями параллелизма.

Для перевода своей полупроводниковой индустрии на выпуск многоядерных процессоров с 2000 г. Intel вложила большие средства в исследования и разработки по этой тематике.

В Intel не считают целесообразным конструировать многоядерные процессоры, состоящие из одного ядра общего назначения и специализированных ядер. Применение универсальных ядер, способных решать самые различные задачи считается более перспективным.

В ПК и серверах будущего будут применяться многоядерные процессоры, поддерживающие такие инновационные технологии, как ускоренный ввод-вывод,

виртуализация (обеспечение виртуальной независимости работы каждого ядра), безопасность и усовершенствованная технология памяти.

3.1. Архитектурные особенности процессора Pentium D

С 2005 года был начат выпуск двухядерных процессоров Pentium D. Двухядерные процессоры Pentium D содержат два независимых ядра на одной кремниевой пластине (рис.3). Каждое ядро имеет собственный кэш второго уровня L2 объемом 1 Мб.

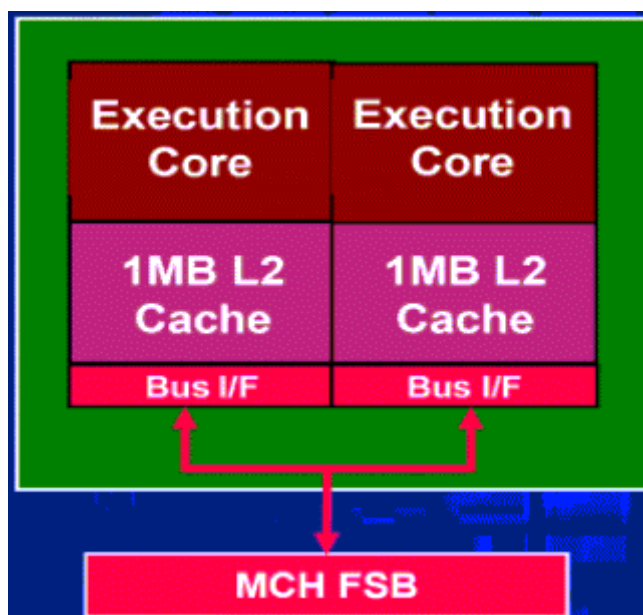


Рис.3. Двухядерный процессор Pentium D

Ядра процессоров базируются на архитектуре NetBurst процессоров Pentium 4. Ядра объединены общей процессорной шиной, работающей на частоте 800 МГц.

Ядра процессоров Pentium D не поддерживают технологию Hyper-Threading. Для двухядерных процессоров она присутствует только в Pentium Extreme Edition, который благодаря этому виден в системе как четырехядерный (рис.4).

Процессоры Pentium D поддерживают 64-битные расширения команд EM64T и технологию XD — защита от атак типа "переполнение буфера". Некоторые модели

Двухъядерный процессор Pentium Extreme Edition был выпущен с тактовой частотой 3,2 ГГц, частотой системной шины 800 МГц и 2 Мб кэш-памяти второго уровня (по 1 Мб на каждое ядро).

Каждое ядро поддерживает Hyper-Threading, поэтому в системе видны четыре процессора.

3.2. Судьба системной шины

Чтобы в полной мере реализовать потенциал роста производительности, обеспечиваемый несколькими ядрами, необходим способ, предоставляющий процессору достаточное количество данных. По мнению экспертов, существующая архитектура системной шины Intel способна удовлетворить требования максимум четырех ядер. В этой архитектуре системная шина связывает центральный процессор с основной памятью. Контроллер памяти, входящий в состав соответствующих наборов микросхем, управляет передачей данных из памяти в центральный процессор и обратно. Многие компании (AMD, Sun Microsystems, IBM) интегрировали контроллеры памяти в кристаллы центральных процессоров. Интегрированный контроллер памяти сокращает уровень задержки (время, требуемое для передачи порции данных от одного компонента системы другому).

Интегрированный контроллер позволил бы Intel увеличить производительность серверных приложений, интенсивно использующих память, и обеспечить доставку данных к нескольким ядрам. Фактически процессор Itanium образца 2007 г. может иметь до восьми ядер. Существует и другой путь, сохранить системную шину за счет добавления к процессорам модулей арбитража шины, которые обеспечат совместное использование процессорами ресурсов, в том числе средств доступа к памяти и систем ввода-вывода. Ряд производителей устанавливает подобные модули в свои многопроцессорные серверы. Во многих серверах с восемью или более процессорами они фактически сгруппированы по четыре. Для Intel не составит большого труда организовать тот же тип архитектуры в процессоре с четырьмя и более ядрами.

20 октября 2006 года Intel продемонстрировала свой первый четырехъядерный процессор для многопроцессорных серверных систем под кодовым названием Tigerton

(на пресс-брифинге в Сан-Франциско был показан в работе четырехпроцессорный сервер, созданный на его основе).

Специализированное высокоскоростное межкомпонентное соединение, соединяющее каждый процессор непосредственно с набором микросхем, обеспечивает более чем двукратный рост производительности и пропускной способности системы.

Чипсет поддерживает модули памяти FB-DIMM (модули DIMM, оснащенные высокоскоростной шиной) и может быть оснащен четырьмя каналами для подключения к модулям памяти, что расширяет полосу пропускания и увеличивает поддерживаемый объем памяти.

Поддержка технологии I/O Acceleration Technology (IOAT), впервые реализована на многопроцессорных платформах.

Четырехядерные процессоры Intel Xeon созданы на базе микроархитектуры Intel Core. Их производительность почти на 50% выше, чем у современного поколения двухядерных процессоров, при сохранении того же уровня энергопотребления.

Двухъядерные процессоры Intel Xeon серии 7000 были изготовлены по 65-нанометровой производственной технологии и были предназначены для установки в серверы с четырьмя или большим числом процессоров. Процессоры этой серии поддерживают технологию Hyper-Threading, благодаря чему способны выполнять 4 вычислительных потока одновременно. Кроме того, они оснащены кэш-памятью третьего уровня объемом 16 МБ, используемой обоими ядрами. Процессоры поддерживают технологии Intel Virtualization Technology и Intel Cache Safe Technology, что позволяет свести к минимуму простои при работе в критических случаях. Двухъядерные процессоры Intel Xeon серии 7100 предназначены для установки в серверы с четырьмя или более процессорами. В состав данного семейства входят процессоры с пониженным энергопотреблением (95 Вт). Процессоры серии 7100 по производительности превосходят процессоры предыдущего поколения почти в два раза, а по показателю «производительность на ватт» – почти в три раза.

3.3. Производительность

Производительность многоядерных процессоров стала зависеть от качества работы программистов. Успех массового программного обеспечения (ПО) зависит от его способности "распараллеливать" задачи. Корпорация Intel взялась за разработку инструментария для написания таких программ и оптимизации существующих.

3.4. Многоядерное будущее

В сентябре 2006 года состоялась демонстрация прототипа процессора Intel, обладающего производительностью в 1 терафлоп (10^{12} операций с плавающей точкой в секунду). Для того чтобы представить, насколько это много, приведем пример.

Первый суперкомпьютер, достигший производительности в 1 терафлоп, ASCI Red, был создан в 1996 году. В его состав входило 4,5 тысячи процессоров Pentium Pro, работавших на частоте 200 МГц.

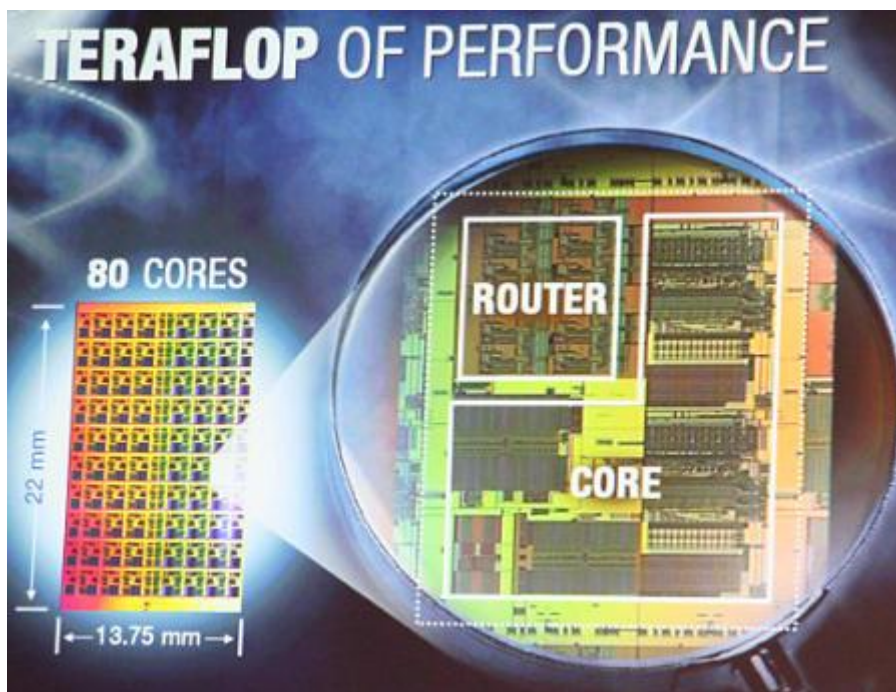


Рис.5. 80-ядерный прототип процессора Intel

Процессор имеет 80 ядер (рис.5) и интегрированную статическую память (SRAM) общим объёмом в 20 Мб (по 256 Кб на ядро). Обмен данными между ядрами процессора осуществляется со скоростью более терабайта в секунду.

Нынешние образцы 80-ядерных процессоров работают на частоте 3,1 ГГц. Ожидается, что коммерческая доступность таких процессоров наступит в течение пяти лет.

В настоящий момент нельзя говорить о том, что перед нами полнофункциональная версия процессоров будущего, скорее - одно из первых воплощений концепта, разработанного в рамках инициативы Intel Tera-Scale, целью которой является доступность для массового покупателя процессоров с производительностью терафлоп уже к 2010-му году.

4. Многоядерные решения AMD

Корпорация AMD (<http://www.amd.com/>) разработала технологию виртуализации Pacifica и технологию безопасности Presidio для серийных многоядерных процессоров. Технология виртуализации Pacifica позволяет запускать на одном компьютере несколько независимых операционных систем и приложений, а технология безопасности Presidio повышает безопасность работы с данными при помощи специальной защищённой области в процессоре.

Многоядерные чипы, спроектированные инженерами AMD, отличаются более низким энергопотреблением, поскольку выполнены на единой подложке, в отличие от кристаллов Intel, где механически соединяются несколько отдельных ядер. Система Cool'n'Quiet автоматически регулирует тактовую частоту и напряжение питания процессора в зависимости от реальной нагрузки, а также скорость вращения процессорного кулера в зависимости от температуры процессора. Конструкция AMD обеспечивает наиболее тесные связи между всеми ядрами, что позволяет не только повысить производительность, но и оптимизировать потребление электроэнергии.

Направление на выпуск многоядерных процессоров заставляет отказаться от схем, оптимальных лишь для одноядерных микросхем. Возможно создание узкоспециализированных моделей процессоров, чётко ориентированных на решение тех или иных задач. К примеру, для серверов производительность процессора в векторных операциях не нужна, поэтому за счёт оптимизации конструкции можно добиться повышенной мощности в поисковых вычислениях и снижения энергопотребления. При этом важно сохранить взаимную совместимость набора инструкций разных модификаций - это не только удешевит производство, но и снимет возможные проблемы при написании программного обеспечения.

Для объединения вычислительных ядер AMD использует архитектуру Direct Connect (прямого соединения), а для связи с набором системной логики используется шина HyperTransport. Третья версия этой шины в три раза превышает пропускную способность HyperTransport 2.0. Совокупная максимальная пропускная способность HyperTransport 2.0. составляет 22,4 Гб/с. В новой версии шины сохраняется её традиционное преимущество - прямая связь между процессором и системами

ввода/вывода, только теперь эти прямые линии связи будут устанавливаться с несколькими ядрами.

Компания летом 2004 г. продемонстрировала оснащенный двумя ядрами 64-разрядный процессор AMD Opteron (с поддержкой системы команд x86). На демонстрации был представлен сервер HP ProLiant DL585 с четырьмя двухъядерными процессорами Opteron, изготовленными по технологии "кремний на изоляторе" (SOI) с соблюдением проектных норм 90 нм. Таким образом, компания AMD заложила фундамент для создания процессоров с двумя ядрами, снабдив свои одноядерные процессоры AMD 64 встроенной инфраструктурой для поддержки второго ядра на той же микросхеме. Двухъядерные процессоры представляют собой естественное расширение технологии AMD 64 с архитектурой прямых соединений.

Компания AMD устранила узкие места в работе внешней шины, характерные для архитектуры x86, она объединила два ядра на одном кристалле вместе с контроллером памяти и подсистемой ввода-вывода (рис.6).

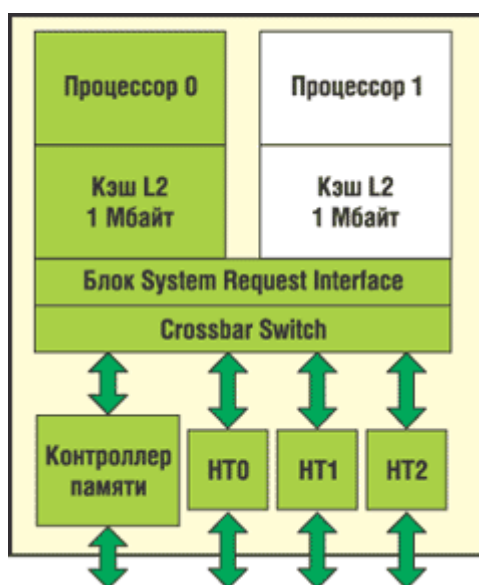


Рис.6. Блок-схема двухъядерного процессора Opteron.

Это позволило улучшить общесистемную производительность и повысить эффективность обработки данных. Архитектура прямых соединений - Direct Connect позволяет подключить несколько процессоров, контроллер памяти и модули ввода-

вывода напрямую. Она нейтрализует недостатки современных системных архитектур и устраняет узкие места при обмене данными. Считается, что данная архитектура сокращает задержки при обращении к памяти, а поскольку средства ввода-вывода непосредственно подключаются к центральному процессору, улучшается баланс между производительностью процессора и пропускной способностью подсистемы ввода-вывода.

В свою очередь, все процессоры соединяются непосредственно друг с другом, что обеспечивает практически линейное повышение производительности для многопроцессорных систем. С ростом числа процессоров пропускная способность памяти растет линейно.

Двухъядерные процессоры AMD 64 поддерживают совместимость с приложениями AMD 64 для платформы x86, что упрощает их массовое распространение.

Следует отметить, что реализация двухъядерности в процессорах AMD несколько отличается от реализации Intel. AMD предлагает несколько иной способ взаимодействия ядер между собой.

Подход Intel заключается в простом помещении на один кристалл двух ядер. При такой организации двухъядерности процессор не имеет никаких специальных механизмов для осуществления взаимодействия между ядрами. Как и в обычных двухпроцессорных системах, ядра общаются посредством системной шины. Соответственно, системная шина разделяется между ядрами процессора и при работе с памятью, что приводит к увеличению задержек при обращении к памяти двух ядер одновременно.

В процессорах AMD дублированы некоторые ресурсы. Хотя каждое из ядер Athlon 64 X2 обладает собственным набором исполнительных устройств и выделенной кэш-памятью второго уровня, контроллер памяти и контроллер шины Hyper-Transport на оба ядра общий. Взаимодействие каждого из ядер с разделяемыми ресурсами осуществляется посредством специального Crossbar-переключателя и очереди системных запросов (System Request Queue). На этом же уровне организовано и взаимодействие ядер между собой, благодаря чему вопросы когерентности кэш памяти решаются без дополнительной нагрузки на системную шину и шину памяти (рис.7)

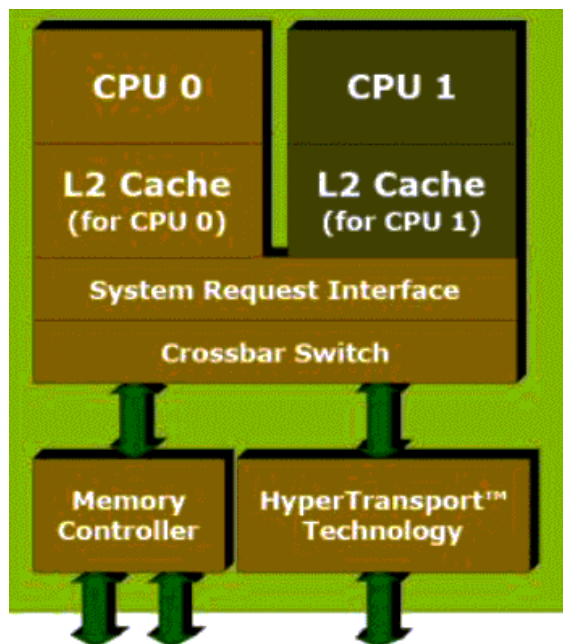


Рис.7. Блок-схема двухъядерного процессора Athlon 64 X2.

Таким образом, единственное узкое место, имеющееся в архитектуре Athlon 64X2 – это пропускная способность подсистемы памяти (6.4 Гбайт в секунду), которая делится между процессорными ядрами.

AMD стремилась сохранить совместимость Athlon 64 X2 с существующими платформами. В результате, эти процессоры стало возможно использовать на тех же самых материнских платах, что и одноядерные Athlon 64. Поэтому, Athlon 64 X2 имеют такой же корпус, двухканальный контроллер памяти с поддержкой DDR400 SDRAM и работают с шиной HyperTransport с частотой до 1 ГГц. Благодаря этому единственное, что потребовалось для поддержки двухъядерных CPU от AMD материнскими платами, – это обновление BIOS. Инженерам AMD удалось вписать в ранее установленные рамки и энергопотребление процессора Athlon 64 X2.

Athlon 64 X2 поддерживают набор инструкций SSE3, а также имеют усовершенствованный контроллер памяти. Среди особенностей контроллера памяти Athlon 64 X2 следует упомянуть возможность использования разномастных модулей DIMM в различных каналах (вплоть до установки в оба канала памяти модулей разного

объёма) и возможность работы с четырьмя двухсторонними модулями DIMM в режиме DDR400.

Процессоры Athlon 64 X2 (Toledo), содержащие два ядра с кэш-памятью второго уровня по 1 Мбайту на каждое ядро, состоят из примерно 233.2 млн. транзисторов и имеют площадь около 199 кв. мм. Кристалл и сложность двухъядерного процессора оказывается примерно вдвое больше кристалла соответствующего одноядерного.

Наличие посредника между процессором и памятью в лице северного моста для многоядерных процессоров Intel увеличивает задержки. При этом пропускная способность памяти не растёт с числом процессоров, а наоборот, процессоры разделяют пропускную способность системы процессор - память (до 6,4 Гбайт/с) между собой. Для компенсации проблемы пропускной способности оперативной памяти в Intel предлагают увеличить ёмкость кэш памяти.

Использование высокоскоростного интерфейса HyperTransport в Opteron/Athlon64 предполагает, как соединение процессоров через HyperTransport с поддержанием когерентности кэша в многопроцессорных серверах, так и прямое — без северного и южного мостов — подсоединение через HyperTransport, мосты для шин PCI-X/PCI-Express, что также повышает производительность. Суммарная пропускная способность ввода/вывода для 8-процессорных систем на базе Opteron 8xx достигает 30,4 Гбайт/с.

Построение серверов на базе Opteron отвечает архитектуре кэш когерентной неоднородной памяти (ccNUMA), пока с небольшим количеством (до 8) процессоров. Можно упомянуть и другие преимущества, например, низкие величины задержек при работе с иерархией памяти. Так, задержка Opteron при выборке из кэш памяти данных первого уровня равна трём тактам. Два порта чтения позволяют выполнять одновременно две таких операции.

Пропускная способность памяти лимитирует производительность многих приложений. Пропускная способность важна, например, при решении задач гидро- и аэродинамики, где используются сеточные методы решения уравнений в частных производных. Классический пример — краткосрочный прогноз погоды. Другой иллюстрацией важности проблемы пропускной способности памяти могут служить коды программ моделирования океана или коды вычислительной химии.

5. Многоядерные процессоры компании IBM

5.1. Двухъядерный (dual-core) процессор Power 4

В 2001 г. корпорация IBM (<http://www.ibm.com/>) разработала свой первый универсальный двухъядерный (dual-core) процессор Power 4, предназначенный для серверов IBM eServer.

Устройство Power 4 было уникально для своего времени тем, что в нем в одном корпусе содержалось два 64-разрядных микропроцессора.

Архитектуру кристалла Power 4 отличало несколько современных решений:

- суперскалярная структура,
- внеочередное исполнение команд,
- большая кэш-память на кристалле,
- специализированный порт для основной памяти,
- высокоскоростные соединения для объединения микропроцессоров в системы с архитектурой распределенной разделяемой памяти.

Каждый процессор Power 4 имел два конвейерных блока для работы с 64-разрядными операндами с плавающей точкой, выбирающих на исполнение по пять команд каждый, и два блока для работы с памятью. Процессоры содержали отдельную кэш-память команд и данных 1-го уровня емкостью по 64 Кбайт каждая. Кроме того, имелась разделяемая (общая) кэш-память 2-го уровня на кристалле (емкостью 1,4 Мбайт) и внешняя кэш-память 3-го уровня (емкостью 32 Мбайт).

Наличие кэш-памяти 2-го уровня, разделяемой двумя процессорами кристалла, а также внешними процессорами других кристаллов через магистрали шириной 128 разрядов, работающие на тактовой частоте более 500 МГц (что обеспечивает пропускную способность свыше 10 Гбайт/с) стало одной из отличительных особенностей Power 4.

Физически кэш-память 2-го уровня емкостью около 1,4 Мбайт состояла из нескольких одинаковых блоков, доступ к которым выполнялся через коммутатор с пропускной способностью на уровне 100 Гбайт/с. Протокол когерентности обеспечивал размещение данных в том блоке кэш-памяти, который использовался для размещения

данных последним. Порт кристалла Power 4, предназначенный для подключения кэш-памяти 3-го уровня емкостью до 32 Мбайт, имел 128 разрядов для каждого из двух направлений пересылки данных. Порт функционировал не на полной тактовой частоте процессоров кристалла, но, тем не менее, при передаче обеспечивал пропускную способность на уровне 13-14 Гбайт/с. Теги для адресации кэш-памяти 3-го уровня были расположены внутри кристалла, что ускоряло реализацию протокола когерентности. Скорость передачи данных между кэш-памятью 3-го уровня и основной памятью могла достигать 12,8 Гбайт/с.

Каждый двоянный процессор Power 4 был упакован в керамический мультипроцессорный модуль (размером 4,5x4,5 дюйма) вместе с тремя другими кристаллами. Эта базовая строительная единица называлась МСМ (MultiChip Module). Такой модуль, содержащий четыре микросхемы Power 4, объединял фактически восемь процессоров. Многослойный керамический корпус модуля МСМ содержал магистрали, соединяющие микросхемы между собой, а также с модулями кэш-памяти и высокоскоростным коммутатором для связи с удаленными процессорами. Четыре микросхемы Power4, образующие восьмипроцессорную конфигурацию, были расположены в модуле под углом 90° друг относительно друга, что позволяло минимизировать длину шин, соединяющих микросхемы между собой, что важно при работе на высоких тактовых частотах. Шины расширения связывали между собой и модули МСМ. В этих шинах использовалась уникальная технология волновой конвейеризации (wave pipelining), обеспечивающая низкие величины задержек.

Стоит отметить, что наряду с параллелизмом на уровне команд процессор Power 4 использовал и параллелизм на уровне потоков (тредов). Базовая многопоточность использовалась еще в процессорах PowerPC RS IV (S-Star) - но в этом процессоре в каждый момент времени выполнялся только один поток команд. Если при выполнении команды одного потока возникала значительная задержка (например, в кэш-памяти отсутствовали необходимые данные), то процессор переключался на другой поток, что уменьшало его простаивание. Однако оба потока совместно использовали многие системные ресурсы (например, архитектурные регистры), поэтому на переключение между потоками тратилось несколько тактов.

Динамическое выявление параллелизма в Power 4 позволило предотвращать простои процессора при трудно выявляемых статических ситуациях, например, промахе при обращении к кэш-памяти.

Компания IBM первой интегрировала в кристалл два процессорных ядра — в технологии Power 4. При создании этой архитектуры она ориентировалась, прежде всего, на рынок высокопроизводительных серверов и суперкомпьютеров, поддерживающих 32-процессорные симметричные многопроцессорные (SMP) - системы.

Микропроцессор Power 4 создавался для работы как с коммерческими (серверными), так и с научно-техническими приложениями. Разработка Power велась на базе RISC-архитектуры (Reduced Instructions Set Computer).

Предусматривается фиксирование длины команд, а также использование универсальных регистров. Команды упрощаются для того, чтобы они могли выполняться за один такт. Понятно, что эффективность вычислений повышается при оптимизации конвейера команд. В процессе создания Power разработчики решили минимизировать не только время каждого цикла, но и полное время, необходимое для выполнения конкретной задачи, для чего была внедрена технология суперскалярной обработки данных.

Вспомним, что каждое из ядер процессора Power 4 обладает собственным кэш первого уровня для данных и инструкций (соответственно LD 1 и LI 1) и общим кэш второго уровня (L2). Кэш память управляется тремя отдельными, автономно работающими контроллерами, которые подключаются к процессорным ядрам через коммутатор (Core Interface Unit - CIU). За такт контроллеры могут выдавать до 32 байт данных. Общий кэш для нескольких ядер хорошо сказывается на производительности процессора благодаря более быстрому доступу к совместно используемой информации, однако при таком способе организации велика вероятность возникновения конфликтов, многоступенчатая же организация кэш позволяет использовать его наиболее эффективно.

Очевидно, что скорость передачи данных между процессором и памятью существенно влияет на потенциал многоядерных устройств. Коммуникация каждого из Power 4 осуществляется коммутатором CIU посредством двух отдельных 256-

битовых шин для выборки инструкций и загрузки данных, а также отдельной 64-битовой шины для сохранения результатов.

Для поддержки когерентности в кэш - памяти второго уровня L2 применена расширенная версия известного протокола MESI с длиной строки 128 байт, а в L3 используется модель с пятью состояниями когерентности (поддержка осуществляется на уровне 128-байтных секторов).

Отдельный функциональный блок, так называемый контроллер матрицы, отвечает за обмен данными между контроллерами кэш-памяти второго и третьего уровней и за коммуникационные функции архитектуры Power 4. В процессоре широко применяются механизмы прогнозирования ветвлений. Адресное пространство составляет 512 Гбайт.

Отметим, что протокол MESI применяет и AMD, точнее, доработанный MESI-MOESI. В нем процессоры могут использовать данные кэш-памяти друг друга, для чего во время чтения информации происходит проверка кэш-памяти соседей и при обнаружении необходимых сведений считывание происходит прямо оттуда. Причем эти данные в оперативной памяти не сохраняются, а владелец информации делает пометку напротив обновленной строки. Любое изменение строки приводит к устареванию аналогичных строк в кэш других процессоров, а в случае вытеснения обновленных строк (Owner-строк) они записываются в оперативную память.

При разработке процессора Power 4 была поставлена задача оптимизации SMP-конфигурации сервера. Для того чтобы сервер на базе Power 4 был готов к работе в многозадачных средах с высокими требованиями к пропускной способности, и были упакованы в один модуль четыре двудерных процессора, образуя восьмипроцессорную SMP систему. Для соединения процессоров между собой вместо центрального коммутатора используется множество независимых шин с топологией типа "точка—точка".

5.2. Следующий шаг - Power 5

Архитектура следующего представителя этой линейки - процессора Power 5, построена на принципах, примененных в Power 4. Два процессорных ядра на одном

кристалле имеют отдельный кэш L1 для данных и инструкций и общий кэш L2. Кэш памяти второго уровня выполнена в виде трех отдельных блоков, у каждого из которых есть свой контроллер. Ядра могут обращаться независимо друг от друга к любому из трех контроллеров.

Ряд важных нововведений начинается с того, что хотя кэш L3 и располагается вне кристалла, он напрямую связан с кэш уровня L2, что снижает задержки при работе с кэш - памятью и улучшает масштабируемость. Система на базе Power 5 может включать до 64 процессорных конфигураций. В микропроцессоре Power 5 IBM впервые реализовала технологию микроразделов (Micro-Partitioning), позволившую представить каждый физический процессор как несколько (до 10) логических. Micro-Partitioning также предусматривает единую консоль для управления системами любых типов и широкий набор системных сервисов для управления рабочими нагрузками и перераспределения ресурсов, что дает возможность выполнить большой объем работы.

Разработчики интегрировали в кристалл процессора Power 5 не только два процессорных ядра, но и элементы управления памятью и заданиями, ранее реализовавшиеся на внешних элементах. Встроенный в процессор контроллер памяти используется и в технологиях, представленных AMD (Athlon 64 и Opteron). Недостаток такой системы — необходимость заменять процессор для перехода на системную плату с новым типом памяти.

Многопоточность в Power 4 реализуется на аппаратном уровне: каждое из двух ядер выполняет свой поток команд, при задержке выполнения одного потока ресурсы обслуживающего его ядра нельзя передать другому потоку. В Power 5 эта задача решена с помощью многопоточной технологии (SMT).

Одновременная многопоточность SMT, при которой процессорное ядро может запрограммировать порядок параллельного выполнения команд из нескольких потоков, впервые появилась только в Power 5. Как известно, технология SMT позволяет системе динамически подстраиваться - если возникает длительная задержка при выполнении команды одного потока (треда), то все исполнительные блоки временно передаются другому потоку. Теоретически SMT можно реализовать не только для двух, но и для большего числа потоков, однако, это слишком усложнит механизм многопоточности, но не даст серьезного выигрыша в производительности.

В Power 5 каждое процессорное ядро может одновременно обрабатывать два потока команд, т. е. работает как два логических процессора, а сам кристалл - как четыре логических процессора. Команды обоих потоков извлекаются из одной и той же кэш-памяти команд 1-го уровня и вместе загружаются в исполнительные блоки. Теоретически оба потока команд после считывания из кэш-памяти команд должны проходить через конвейер и использовать ресурсы физического процессора без конфликтов между потоками. В этом случае применение SMT удваивает производительность процессора. Однако если между потоками возникнет конфликт из-за ресурсов процессора, то одному из них придется ждать, и прирост производительности оказывается меньше 100%.

SMT организует выполнение команд из потоков следующим образом. Несколько команд из каждого потока можно загрузить в набор конвейеров в течение каждого такта, причем любые пять команд из обоих потоков, представляющие до двух операций загрузки и сохранения, двух операций двоичной арифметики, двух операций с плавающей точкой и ветвление, допустимо загрузить в один конвейер в течение одного такта. За один такт по стадиям конвейера проходит только одна команда из одного из потоков. Физический процессор определяет порядок выполнения команд на основе набора правил с учетом зависимости операндов, времени ожидания обработки команд и необходимости внеочередного выполнения некоторых команд.

Каждый физический процессор содержит кэш-память данных и команд 1-го уровня, которые совместно используются обоими логическими процессорами. Как говорилось выше, кэш-память 2-го уровня у физических процессоров (ядер) Power 5 общая, поэтому она совместно используется четырьмя потоками команд. Для улучшения производительности SMT при выполнении различных приложений в POWER5 реализованы динамическая балансировка ресурсов и подстраиваемый приоритет потоков.

Распределение приоритетов выполнения вычислительных потоков происходит на аппаратном уровне. Микропроцессор поддерживает восемь уровней приоритета (от седьмого, соответствующего наивысшему приоритету потока, до нулевого). Поток с более высоким приоритетом предоставляется большая часть ресурсов. Когда

приложению требуется использовать всю полосу пропускания для обмена данными с памятью, процессор автоматически переходит в однопоточный режим.

Очевидно, что применение SMT позволяет каждому ядру выполнять больше команд за такт, вследствие чего увеличивается и энергопотребление.

Решение этой проблемы обеспечивают два механизма:

- при низком приоритете обоих потоков (0 или 1) включается встроенная технология энергосбережения;
- возможно временное отключение тех компонентов процессора, которые не потребуются во время следующих тактов.

В результате, при одинаковых затратах энергии, Power 5 выполняет на 50% больше инструкций, чем Power 4.

Технология сверхплотной компоновки, особенность которой заключается в применении общей оперативной памяти и межузловых соединений с большой пропускной способностью, позволяет реализовать высокоскоростные соединения между восемью процессорами Power 5. Четыре кристалла Power 5 с четырьмя кристаллами кэш - памяти L3 упаковываются в многокристальный модуль Multichip Module (MCM) размерами 95*95 мм. Так как Power 5 нацелен на работу с системами хранения данных нового поколения, то поддерживается адресация хранилища данных емкостью до 96 Пентабайт! Отметим также, что Power 5 обратно совместим с Power 4 на уровне программного обеспечения.

5.3. Power 5+

Следующим шагом для IBM стал Power 5+ (рис. 8), представляющий собой так называемый "сервер-на-кристалле". Он содержит два процессора, поддерживающих SMT-технология, высокопроизводительный системный коммутатор, встроенную кэш-память объемом до 72 Мбайт, интерфейс ввода-вывода. Пиковая пропускная способность шины процессор—память составляет 42,6 Гбайт/с. Серверы, оснащенные микропроцессорами Power 5+, сконструированы с учетом требований к вычислительным системам малого и среднего бизнеса, они учитывают ограниченность ресурсов, характерную для многих таких компаний.

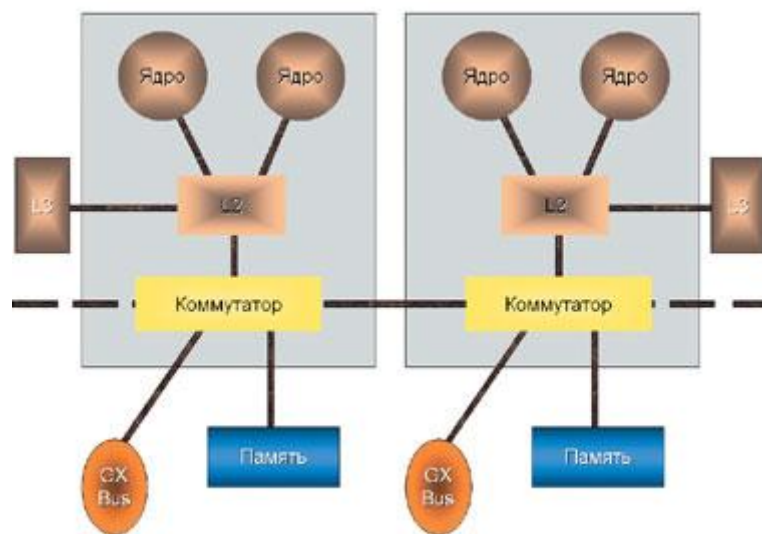


Рис.8. Схема процессора Power 5+

У них есть удобный программный инструментарий, и они обладают возможностью "внутреннего масштабирования" за счет реализации технологий виртуализации Micro-Partitioning. Использование логических разделов позволяет создавать ИТ-системы уменьшенных размеров и повышенной производительности.

5.4. Power PC 970MP

Технология Power была положена и в основу процессоров Power PC. Последний продукт этой серии, Power PC 970MP, представляет собой первую двуядерную 64-битовую версию Power PC, созданную на базе архитектуры Power 4 с использованием 90-нм технологии. Между 64- и 32-битовыми конфигурациями возможно динамическое переключение. Эффективность работы поддерживается механизмом SMP-оптимизации. Каждое ядро в PowerPC 970MP обладает собственной кэш-памятью L2, а также независимым термодиодом и шиной питания. По сравнению с предыдущими модификациями у него увеличена пропускная способность процессорной шины и размер кэш-памяти уровня L2. Раздельный кэш L2 позволяет отключать или переводить в спящий режим одно из ядер. За такт может быть выполнено до восьми

инструкций. Для сравнения: процессоры AMD на базе архитектуры K8 выполняют не больше шести инструкций за такт. Тактовые частоты ядер регулируются синхронно.

На игровых приставках Xbox 360 вместо процессора Intel, установленного в первой приставке Xbox, теперь применяется кристалл на базе PowerPC, который IBM назвала "компьютер-на-кристалле" (SoC или System-on-Chip).

На одной подложке интегрированы три идентичных ядра, объединенных диспетчерской шиной XBAR, причем каждое из ядер способно выполнять по два независимых потока команд. Этот процессор, созданный специально для игровых и мультимедийных приложений, эффективен и для обработки больших объемов потоковых данных. В процессоре реализована также возможность обхода кэш-памяти второго уровня L2 с пересылкой до восьми пакетов данных предварительной выборки непосредственно в кэш L1 каждого из трех ядер процессора. При этом части кэш памяти L2 могут быть распределены в качестве буферов для потоковых данных. Xbox 360 стала первой приставкой с процессором седьмого поколения, появившейся на рынке.

5.5. Процессорный элемент Cell -"суперкомпьютер-на-кристалле"

На базе архитектуры Power в рамках трехстороннего партнерского соглашения между IBM, Sony и Toshiba Corporation был создан и процессорный элемент Cell, представляющий собой "суперкомпьютер-на-кристалле". Его архитектура включает восемь взаимодополняющих вычислительных элементов SPE (Synergistic Processor Element) и ядро на базе Power. Все эти компоненты соединены между собой скоростной шиной EIB.

Встроенный двухканальный контроллер памяти способен работать с памятью XDR компании Rambus (максимальный объем 256 Мбайт). Связь с остальными компонентами производится посредством системной шины FlexIO с пиковой тактовой частотой 6,4 ГГц, при этом FlexIO может связываться с данным элементом и другие процессоры Cell.

Специализированные микрокомпьютеры SPE рассчитаны на работу на тактовых частотах свыше 4 ГГц, поддерживают массовую обработку данных с плавающей

запятой и поддержку нескольких ОС одновременно. Отдельно взятый вычислительный элемент Cell имеет теоретическую мощность 250 GFLOPS (миллиарды операций с плавающей запятой в секунду). Кроме того, процессор оптимизирован для выполнения широкополосных медиаприложений (например, игры, развлекательные программы, видео) и обладает встроенной технологией управления энергопотреблением. На базе элемента Cell стали выпускаться не только игровые приставки, но и серверы, так как он хорошо справляется с серьезными вычислительными задачами.

6. Многоядерные процессоры SPARC компании Sun Microsystems

В начале 2004 г. Sun Microsystems (<http://www.sun.com/>) выпустила оснащенный двумя ядрами процессор UltraSparc IV для серверов Sun Fire V и представила проект процессора нового поколения - UltraSPARC IV+ (Panther).

Он выполнен на базе технологии CMT (Chip Multithreading) - многопоточность на кристалле, и знаменует собой очередной шаг в реализации стратегии Throughput Computing ("высокая пропускная способность"). UltraSPARC IV+ был разработан на базе 90-нм производственной технологии Texas Instruments. По сравнению с процессором UltraSPARC IV он позволяет удвоить производительность приложений благодаря увеличению объема кэш-памяти и буферов, улучшенному механизму прогнозирования ветвления, расширенным возможностям упреждающей выборки из памяти и новым вычислительным возможностям. Кроме того, в UltraSPARC IV+ применяется трехуровневая иерархия кэш-памяти, включающая интегрированную на кристалле кэш-память 2-го уровня объемом 2 Мбайт и внешнюю кэш-память 3-го уровня объемом 32 Мбайт.

Подобно UltraSPARC IV, новый процессор имеет два ядра, интегрированных на одном кристалле. Кроме новых функций увеличения производительности, процессор UltraSPARC IV+ имеет значительно более высокую тактовую частоту (первоначально 1,8 ГГц), что обеспечивает наивысшую пропускную способность по сравнению с другими моделями процессоров UltraSPARC. По сравнению с UltraSPARC IV производительность каждого потока выросла примерно вдвое.

Так же, как процессор UltraSPARC IV, UltraSPARC IV+ обеспечивает бинарную совместимость с предыдущими поколениями процессоров архитектуры SPARC. Благодаря этому пользователи сохраняют средства разработки и прикладное ПО.

Кроме того, пользователям легко модернизировать существующие системы, повысив их производительность и надежность, - процессоры могут устанавливаться в существующие системы семейства Sun Fire и работать вместе с уже установленными в системе процессорами UltraSPARC IV и UltraSPARC III.

6.1. Проект UltraSPARC T1 (Niagara)

14 ноября 2005 года компания представила новый многоядерный процессор UltraSPARC T1, ранее известный под кодовым названием Niagara.

Проект Niagara - главная часть работы Sun, нацеленной на сохранение актуальности семейства процессоров SPARC на фоне широкой популярности архитектуры x86 от Intel и AMD и растущей мощности процессоров POWER от IBM. Выпуск кристаллов Niagara планировался на 2006 г., как еще одна демонстрация преимуществ технологии CMT. В 2008 г. планируется выпустить второе поколение процессора Niagara II и микросхемы с кодовым названием Rock. Основные технические характеристики UltraSPARC T1:

Архитектура SPARC девятого поколения

- Кэш-память второго уровня на кристалле
- Поддержка шифрования с открытым ключом RSA
- 48-битная виртуальная и 40-битная физическая адресация
- Поддержка страниц четырех размеров: 8 Кбайт, 64 Кбайта, 4 Мбайта и 256 Мбайт
- Поддержка технологии Hypervisor

На рисунке 9 приведена схема обработки 4-х потоков в одном ядре для 4-х ядерного процессора.

Версии процессоров с 4, 6 или 8 ядрами:

- Первичная кэш-память инструкций объемом 16 Кбайт на каждое ядро
- Первичная кэш-память данных объемом 8 Кбайт на каждое ядро
- Единая кэш-память второго уровня объемом 3 Мбайта
- До восьми ядер, 4 потока в каждом ядре
- Четыре 144-битных контроллера оперативной памяти DDR2 533 SDRAM ECC
- До четырех модулей оперативной памяти на каждый контроллер (до 16 DIMM)
- Двухканальный режим оперативной памяти

- Интерфейс JBUS с пиковой эффективной пропускной способностью 3,1 Гбайт/с, 128-битная шина, частота - от 150 до 200 МГц
- 0,09-микронный технологический процесс, КМОП
- Тактовая частота - 1,0 или 1,2 ГГц
- Энергопотребление: 72 Вт, пиковая - 79 Вт

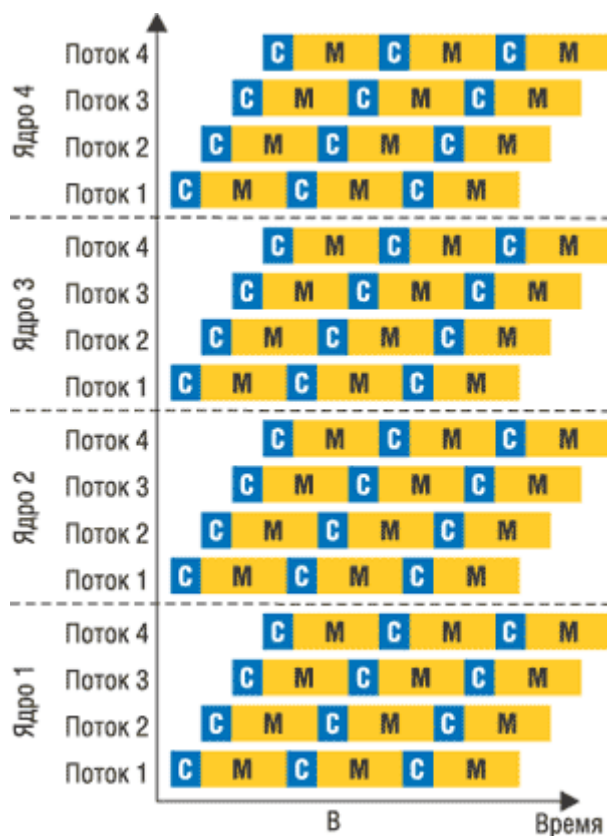


Рис. 9. Обработка одновременно 4-х потоков в одном ядре.

Процессор UltraSPARC T1 состоит из восьми ядер, каждое из которых способно параллельно обрабатывать четыре потока, так что каждый чип может работать с 32 параллельными потоками, что чрезвычайно важно для таких типов приложений, как базы данных или поиск в интернете. Разработчики утверждают, что использованные технологии позволяют вдвое сократить число веб-серверов во всем мире, в результате чего удастся добиться снижения выбросов диоксида углерода от сжигания нефти и угля на объем, эквивалентный посадке миллиона акров деревьев. Экономия на

электроэнергии, по расчетам Sun, составит порядка 14 миллиардов долларов США в год.

Изначально низкое энергопотребление достигнуто в UltraSPARC T1 весьма тривиальным способом - невысокой тактовой частотой, которая составляет всего 1,2 ГГц. Вместе с тем и архитектура процессора претерпела изменения: почти вся кэш-память сервера, системная память и элементы ввода/вывода встроены непосредственно в сам процессор. По существу, UltraSPARC T1 представляет собой один из первых в мире полноценных серверов на одном чипе. Помимо снижения энергопотребления и тепловыделения, такая архитектура существенно упрощает конструирование самих серверов и позволяет достичь высочайшего уровня производительности при сохранении компактных размеров.

Многие микропроцессоры в состоянии простоя потребляют лишь чуть меньше энергии, чем при работе с максимальной нагрузкой, поэтому тепловыделение этих чипов практически не снижается и электроэнергия расходуется неэффективно. Поэтому, перед разработчиками была поставлена задача, изменить эту ситуацию, предусмотрев значительное снижение энергопотребления в режиме простоя, что позволило бы более эффективно охлаждать кристаллы.

В результате, серверы на основе UltraSPARC T1 не требуют большого количества массивных и громоздких систем охлаждения, как серверы на процессорах традиционной конструкции. По некоторым данным, порядка 80 процентов всех современных центров обработки и хранения данных не могут расширяться из-за проблем, связанных с электропитанием и охлаждением, а порядка 40 процентов дата-центров сталкиваются с необходимостью усовершенствования систем охлаждения. Эксперты отмечают, что за последние пять лет бурного развития интернета, тепловыделение оборудования центров обработки данных выросло в четыре раза.

Производитель планировал процессор UltraSPARC T1, как чип, рассчитанный именно на дата-центры, работающие как у провайдеров телекоммуникационных и прочих услуг, так и в областях производства, финансов, розничной торговли, энергетики, образования и здравоохранения. Среди приложений, на которые рассчитан процессор - объединенные веб-серверы, серверы Java-приложений и виртуальные Java-машины, инфраструктурные сервисы, а также на деловые системы ERP и CRM.


Первыми серверами на основе UltraSparc T1 стали две модели: 1,75-дюймовая T1000 с кодовым названием Erie и 3,5-дюймовая T2000 с кодовым названием Ontario.

7. Заключение

С появлением многоядерных процессоров открывается новая эра "настольных вычислений". Новые технологии позволяют работать в многозадачных средах с одновременным выполнением нескольких активных и фоновых приложений, повысить эффективность и снизить энергопотребление при одновременном запуске множества приложений, увеличить количество пользователей, работающих одновременно на одном ПК. Многоядерные процессоры обеспечивают более высокую производительность для потоковых приложений.

Список источников

1. Шагурин И.И. Pentium 4 — новая ступень развития микропроцессорной техники // Chip News. — 2000. — № 9. — С. 18–20.
2. Технология Hyper-Threading (HT) компании Intel («Мир ПК», №2/02, с. 10).
3. М. Кузьминский, «64-разрядные микропроцессоры AMD». Открытые системы, № 4, 2002.М.
4. Алексей Борзенко. Многоядерные процессоры. Платформы и технологии | №5/2005
5. Сергей Орлов. Многоядерность, параллелизм, виртуализация. Журнал «LAN», #05, 2006 год // Издательство «Открытые системы»
6. Intel®Pentium® 4. Особенности архитектуры процессоров Pentium 4
<http://www.iu4.bmstu.ru>; <http://cisc.narod.ru/INTEL/intel9.htm>;
<http://fait.org.ru/club/htm/>
7. Станислав Гарматюк. Pentium 4 - Взгляд вблизи.
<http://www.pkinform.ru/newspaper/2002/17/p4.html>
8. Василий Леонов. Двухъядерные процессоры Intel: выбираем лучший
http://www.citforum.ru/hardware/microcon/intel_choice/index.shtml
9. Алексей Борзенко. Многоядерные процессоры.
<http://www.linknews.ru/news22485.html>
10. IDF Fall`06: Intel демонстрирует 80-ядерный процессор с производительностью в 1 терафлоп. 28 Сентября, 2006. <http://www.ixbt.com/news/all/index.shtml>
11. Многоядерные процессоры: первые попытки.
http://www.compdoc.ru/comp/cpu/multikernel_cpu_first_attempts/ Компьютерная документация от А до Я
12. Сетевые решения: Многоядерные процессора Intel для серверов и рабочих станций. <http://www.nestor.minsk.by/sr/news/2006/10/2401.html>
13. Дмитрий Чеканов. 1 января 2006. Начало 65-нм эпохи. 65-нм процессор Intel Pentium Extreme Edition 955: новая надежда?
<http://www.thg.ru/cpu/20060101/onepage.html>

14. Олег Нечай. Многоядерные планы AMD: секретность, и ещё раз секретность.
Опубликовано 16 марта 2006 года. <http://www.terralab.ru/system/257824/>
15. Марина Старкова Многоядерность RISC-процессоров
10.07.2006 <http://www.morepc.ru/processor/prc180920061.html?print>
16. Процессор UltraSPARC T1 - тяжелая артиллерия Sun. 20 ноября 2005.
CITnews_ru.files/main.css (ссылка на источник:  <http://www.terralab.ru>)
17. Многоядерность RISC-процессоров. Полный вариант статьи см. на «Мир ПК-диске». http://www.MorePC_Ru.htm
18. Процессор UltraSPARC T1 - тяжелая артиллерия Sun. 20 ноября 2005.
www.citnews.ru/20-11-2005/
19. Автор: Gavric. 20.07.2006 Большое тестирование: современные двухъядерные процессоры для настольных компьютеров.
<http://www.fccenter.ru/online.shtml?articles/hardware/processors/18441>
20. Двухъядерные процессоры Intel и AMD: теория. Часть 1
Страница 2. Многоядерные процессоры
<http://www.cyberguru.ru/hardware/processors/intel-2core-1-page2.html>
21. Эволюция многоядерной процессорной архитектуры Intel Core: Conroe, Kentsfield, далее по расписанию.
22. Владимир Романченко. Стр.1 - Новая архитектура Intel Core Дата: 27.06.2006. http://www.3dnews.ru/cpu/new_core_conroe
Стр.2 - Особенности Intel Core - I
http://www.3dnews.ru/cpu/new_core_conroe/index2.htm
Стр.3 - Особенности Intel Core - II.
http://www.3dnews.ru/cpu/new_core_conroe/index3.htm
23. Вячеслав Кононов. 12.02.2007. 80-ядерный процессор от Intel: новые подробности.
http://www.3dnews.ru/news/80_yadernii_protsektor_ot_intel_novie_podrobnosti
- ISSCC 2007: 6-ГГц Cell и 98-Вт 80-ядерник от Intel;
- IDF Fall 2006: 45 нанометров, 80 ядер и один терафлоп.
24. ISSCC 2007: 6-ГГц Cell и 98-Вт 80-ядерник от Intel
http://www.3dnews.ru/news/isscc2007_6ggz_cell_i_98vt_80yadernik_ot_intel

25. IDF Fall 2006: 45 нанометров, 80 ядер и один терафлоп

http://www.3dnews.ru/news/idf_fall_2006_45_nanometrov_80_yader_i_odin_teraflop