**Архитектура и производительность серверных ЦП**

Сегодня центральные процессоры (ЦП) — неотъемлемая часть современной компьютерной техники, но так было не всегда. Первые «серверы» в современном понимании (мэйнфреймы) не имели ЦП как таковых. Вычислительные функции в них выполняли неспециализированные наборы системной логики (стандартные микросхемы и прочие полупроводниковые компоненты). Однако по мере развития компьютерных технологий появилась возможность создавать все более сложные микросхемы, в том числе специализированные — процессоры.

Когда появились первые персональные компьютеры и начали быстро расти их продажи (в 1970-1980-х годах), рынок серверного оборудования серьезно изменился: наметились тенденции взаимопроникновения технологий и архитектур из сегмента серверной техники в сегмент персональных компьютеров, и наоборот. Не стали исключением и ЦП: некоторые серверные модели после незначительной модернизации стали использоваться в персональных компьютерах, а ЦП, изначально разработанные для массового рынка, проникали на рынок серверов и суперкомпьютеров. Наиболее очевидная причина этого — экономический фактор: параллельная разработка двух или более процессорных архитектур может оказаться непосильной даже для довольно крупной компании, а конечная стоимость обоих решений будет значительной, что неминуемо ухудшит конкурентные возможности конечной продукции на рынке. Хотя процессорные архитектуры, ориентированные исключительно на серверы и суперкомпьютеры, до сих пор существуют, более распространены «универсальные» процессорные ядра, пригодные как для массового, так и для серверного рынка.

**Общая архитектура ЦП**

**Функциональные устройства ЦП**

Любой ЦП независимо от его предназначения располагает рядом функциональных устройств (ФУ): блоками целочисленных вычислений (Integer Execution Unit, E-box, арифметико-логическое устройство), обработки адресов (Address Translation Unit, A-box), контроллером системной шины (System Bus Controller, B-box), декодером команд (Instruction Decoding Unit, I-box). Для проведения вычислений и хранения промежуточных результатов используется так называемый файл регистров (их может быть несколько), состоящий из фиксированного количества записей определенной длины (регистров). Практически во всех современных ЦП реализованы и средства предсказания направления условных переходов. С начала 1990-х гг. общепринятой стала практика внедрения в ядро ЦП блока вещественных вычислений (Floating-Point Unit, F-box). Если дизайн определенной аппаратной платформы предполагал использование кэш-памяти с прямым доступом, то в ЦП вводился блок взаимодействия с кэш-памятью (Cache Memory Interface Unit, C-box). В современных ЦП нередко также реализованы контроллеры памяти, периферийных шин и т. д.

**Набор команд ЦП**

В зависимости от представления команд процессоры принято делить на RISC (Reduced Instruction Set Computing, вычисления при помощи ограниченного набора команд) и CISC (Complex Instruction Set Computing, вычисления при помощи расширенного набора команд). В настоящее время «чистокровных» представителей того или иного семейства встретить все труднее — у большинства ЦП есть черты обеих архитектур.

Архитектура RISC появилась как попытка избавиться от недостатков, присущих CISC-процессорам. В 1970-х гг. были проведены исследования, которые показали, что около 20% всех команд CISC-процессора занимают при выполнении около 80% всего процессорного времени, а остальные 80% команд — оставшиеся 20%, т. с. налицо явный дисбаланс — некоторые команды остаются практически невостребованными. В то же время каждая дополнительная команда в наборе увеличивает стоимость ЦП, поскольку для ее выполнения в процессоре должны быть соответствующие исполнительные блоки. Возникла идея свести набор команд к минимуму, зафиксировать их длину для удобства декодирования, а непосредственный доступ к памяти (загрузку данных в регистры ЦП из памяти и копирование из регистров в память) разрешить лишь некоторым из них. Это позволило получить компактное ядро ЦП, способное работать на более высоких частотах при неизменных проектных нормах.

Естественно, что и у RISC, и у CISC есть свои достоинства. RISC обеспечивает эффективную суперскалярность (параллельность) вычислений функциональными устройствами ЦП и большее количество выполненных команд в расчете на один такт ЦП. CISC предполагает более эффективную работу ЦП при небольшом числе функциональных устройств и наличии не более чем одного конвейера в расчете на функциональное устройство (при увеличении количества конвейеров прирост производительности оказывается небольшим из-за ограничений, налагаемых декодером команд).

**ЦП архитектуры х86: CISC или RISC?**

Процессоры архитектуры х86 изначально имели CISC-архитектуру, но со временем стали приобретать свойства RISC. Ядра этих ЦП претерпели кардинальные изменения: они стали проектироваться во многом по принципам RISC, но в то же время для сохранения совместимости с существующим программным обеспечением снаружи в ЦП все так же поступали обычные команды CISC. Для того чтобы такая модель ЦП оказалась работоспособной, требуется обязательное внутреннее преобразование CISC-команд в одну или несколько RISC-команд, непосредственно выполняющихся ядром ЦП.

Такой подход помог улучшить суперскалярность вычислений (когда команды разного рода одновременно обрабатываются на разных конвейерах ФУ) и облегчить их конвейеризацию в целом, но не избавил архитектуру от всех недостатков, главный из которых — неудобство параллельного декодирования команд CISC, имеющих разную длину: невозможно определить начало следующей команды до завершения декодирования предыдущей.

Первым из ЦП архитектуры х86, спроектированным с RISC-ядром, стал уже забытый Nx586 компании NexGen, представленный в марте 1994 г. В ноябре 1995 г. появился Pentium Pro компании Intel, а в марте 1996-го — AMD К5, построенный на базе модернизированной фирменной архитектуры 29К.

CISC-процессоры архитектуры х86 еще некоторое время выпускались, преимущественно Intel (Pentium и Pentium MMX) и Cyrix (6x86 и 6х86МХ/6х86МП), но дни их были сочтены. В последующие годы архитектура Intel P6, в основу которой лег вышеупомянутый Pentium Pro, продолжала экстенсивно развиваться, а позже ее сменила архитектура NetBurst.

В апреле 1997 г. AMD выпустила Кб, основанный на RISC-ядре Nx686 (NexGen была поглощена AMD в начале 1996 г.), и развивала этот модельный ряд до 2000 г., невзирая на отставание в производительности от архитектуры Intel P6. В апреле 1999 г. AMD представила новый ЦП К7, получивший официальное название Athlon, и успешно продолжает развивать его архитектуру и по сей день — ядро К8 (Athlon 64, Athlon 64 FX, Opteron) представляет собой эволюционное развитие ядра К7. Подводя итоги вышесказанному, можно сделать вывод, что ЦП архитектуры х86 перестали быть CISC, но и RISC в полном понимании не стали (подобный переход не обеспечит сохранения обратной совместимости с существующим программным обеспечением), поэтому их можно классифицировать как псевдо-RISC.

**Концепция VLIW**

Интересная разновидность подхода RISC — концепция VLIW (Very Long Instruction Word), известная также как EPIC (Explicitly Parallel Instruction Computing, вычисления с явной параллельностью). Ее суть — эффективное распределение нагрузки методом «примитивный ЦП — умный компилятор»: ЦП располагает большим количеством конвейеров функциональных устройств, работающих параллельно, а всю работу по формированию эффективного потока команд выполняет компилятор. Соответственно, алгоритмы и блоки внеочередного выполнения (out-of-order execution) не используются совсем. Декодер ЦП производит выборку из потока «связок» команд, а планировщик направляет их в очереди на нужные функциональные устройства.

Наиболее известные представители этого подхода — процессоры Crusoe и Transmeta Efficeon, а также Intel Itanium и Itanium 2.

**Кэш-память в ЦП**

**Статическая и динамическая память**

В каждом современном ЦП предусмотрено некоторое количество статической памяти, работающей на частоте ядра. Именно статической, поскольку использование динамической памяти в этих целях представляется крайне нерациональным.

Одна ячейка статической памяти состоит из шести транзисторов и двух резисторов (для техпроцессов с проектными нормами до 0,5 мкм могли быть использованы только четыре транзистора на одну ячейку, с дополнительным слоем поликремния и с более жесткими ограничениями по максимальной тактовой частоте), в то время как аналогичная структура динамической памяти состоит из одного транзистора и одного конденсатора. Быстродействие статической памяти намного выше (емкость, используемая в динамической памяти, имеет определенную скорость зарядки до требуемого уровня, определяющую «частотный потолок»), но из-за большего количества составляющих элементов она существенно дороже в производстве и отличается более высоким энергопотреблением. Битовое значение ячейки статической памяти характеризуется состоянием затворов транзисторов, а динамической — уровнем заряда конденсатора. Так как конденсаторы имеют свойство с течением времени разряжаться, то для поддержания их состояния требуется регулярная перезарядка (для современных микросхем динамической памяти — приблизительно 15 раз в секунду). Кроме того, при операции чтения из ячейки динамической памяти конденсатор разряжается, т. е. ячейка утрачивает свой первоначальный заряд, а следовательно должна быть перезаряжена.

Очевидно, что статическая память используется там, где требуется максимальное быстродействие (регистры ЦП, кэш-память, локальная память сигнального процессора), а динамическая — там, где приоритетом является объем, а не скорость чтения-записи (оперативная память, буферы периферийных устройств).

**Причины внедрения кэш-памяти**

Явная необходимость в кэш-памяти при проектировании массовых ЦП проявилась в начале 1990-х гг., когда тактовые частоты ЦП значительно превысили частоты системных шин, и, в частности, шины памяти. В настоящее время частоты серверных ЦП достигают почти 4 ГГц, а оперативной памяти, массово применяемой в серверах, — только 400 МГц (200 МГц с удвоением благодаря передаче по обоим фронтам сигнала). В этой ситуации при прямом обращении к памяти функциональные устройства ЦП значительную часть времени простаивают, ожидая доставки данных. В какой-то мере проблемы быстродействия оперативной памяти могут быть решены увеличением разрядности шины памяти, но даже в серверах младшего уровня нередко встречается 8-16 гнезд для модулей памяти, поэтому такое решение усложняет дизайн системной платы. Проложить же 256- или даже 512-бит шину к расположенной внутри кристалла ЦП кэш-памяти сравнительно несложно. Таким образом, эффективной альтернативы кэш-памяти в современных высокопроизводительных системах не существует.

**Иерархическая модель кэш-памяти**

Как правило, кэш-память имеет многоуровневую архитектуру. Например, в компьютере с 32 Кбайт внутренней (в ядре ЦП) и 1 Мбайт внешней (в корпусе ЦП или на системной плате) кэш-памяти первая будет считаться кэш-памятью 1-го уровня (L1), а вторая — кэш-памятью 2-го уровня (L2). В современных серверных системах количество уровней кэш-памяти может доходить до четырех, хотя наиболее часто используется двух- или трехуровневая схема.

В некоторых процессорных архитектурах кэш-память 1-го уровня разделена на кэш команд (Irstruction Cache, I-cache) и кэш данных (Data Cache, D-cache), причем необязательно одинаковых размеров. С точки зрения схемотехники проще и дешевле проектировать раздельные I-cache и D-cache: выборку команд проводит 1-bох, а выборку данных — Е-box и F-box, хотя в обоих случаях задействуются А-box и С-box. Все эти блоки велики, и обеспечить им одновременный и быстрый доступ к одному кэшу проблематично. Кроме того, это неизбежно потребовало бы увеличения количества портов доступа, что также усложняет задачу проектирования.

Так как I-cache и D-cache должны обеспечивать очень низкие задержки при доступе (это справедливо для любого кэша L1), приходится жертвовать их объемом — обычно он составляет от 16 до 32 Кбайт. Ведь чем меньше размер кэша, тем легче добиться низких задержек при доступе.

Кэш-память 2-го уровня, как правило, унифицирована, т. е. может содержать как команды, так и данные. Если она встроена в ядро ЦП, то говорят о S-cache (Secondary Cache, вторичный кэш), в противном случае — о B-cache (Backup Cache, резервный кэш). В современных серверных ЦП объем S-cache составляет от одного до нескольких мегабайт, a B-cache — до 64 Мбайт. Если дизайн ЦП предусматривает наличие встроенной кэш-памяти 3-го уровня, то ее именуют T-cache (Ternary Cache, третичный кэш). Как правило, каждый последующий уровень кэш-памяти медленнее, но больше предыдущего по объему. Если в системе присутствует B-cache (как последний уровень модели кэш-памяти), то он может контролироваться как ЦП, так и набором системной логики.

Если в момент выполнения некоторой команды в регистрах не окажется данных для нее, то они будут затребованы из ближайшего уровня кэш-памяти, т. е. из D-cache. В случае их отсутствия в D-Cache запрос направляется в S-cache и т. д. В худшем случае данные будут доставлены непосредственно из памяти. Впрочем, возможен и еще более печальный вариант, когда подсистема управления виртуальной памятью операционной системы (ОС) успевает вытеснить их в файл подкачки на жесткий диск. В случае доставки из оперативной памяти потери времени на получение нужных данных могут составлять от десятков до сотен тактов ЦП, а в случае нахождения данных на жестком диске речь уже может идти о миллионах тактов.

**Ассоциативность кэш-памяти**

Одна из фундаментальных характеристик кэш-памяти — уровень ассоциативности — отображает ее логическую сегментацию. Дело в том, что последовательный перебор всех строк кэша в поисках необходимых данных потребовал бы десятков тактов и свел бы на нет весь выигрыш от использования встроенной в ЦП памяти. Поэтому ячейки ОЗУ жестко привязываются к строкам кэш-памяти (в каждой строке могут быть данные из фиксированного набора адресов), что значительно сокращает время поиска. С каждой ячейкой ОЗУ может быть связано более одной строки кэш-памяти: например, n-канальная ассоциативность (n-way set associative) обозначает, что информация по некоторому адресу оперативной памяти может храниться в n мест кэш-памяти.

Выбор места может проводиться по различным алгоритмам, среди которых чаще всего используются принципы замещения LRU (Least Recently Used, замещается запись, запрошенная в последний раз наиболее давно) и LFU (Least Frequently Used, запись, наименее часто запрашиваемая), хотя существуют и модификации этих принципов. Например, полностью ассоциативная кэшпамять (fully associative), в которой информация, находящаяся по произвольному адресу в оперативной памяти, может быть размещена в произвольной строке. Другой вариант — прямое отображение (direct mapping), при котором информация, которая находится по произвольному адресу в оперативной памяти, может быть размещена только в одном месте кэш-памяти. Естественно, этот вариант обеспечивает наибольшее быстродействие, так как при проверке наличия информации контроллеру придется «заглянуть» лишь в одну строку кэша, но и наименее эффективен, поскольку при записи контроллер не будет выбирать «оптимальное» место. При одинаковом объеме кэша схема с полной ассоциативностью будет наименее быстрой, но наиболее эффективной.

Полностью ассоциативный кэш встречается на практике, но, как правило, у него очень небольшой объем. Например, в ЦП Cyrix 6x86 использовалось 256 байт такого кэша для команд перед унифицированным 16-или 64-Кбайт кэшем L1. Часто полноассоциативную схему применяют при проектировании TLB (о них будет рассказано ниже), кэшей адресов переходов, буферов чтения-записи и т. д. Как правило, уровни ассоциативности I-cache и D-cache довольно низки (до четырех каналов) — их увеличение нецелесообразно, поскольку приводит к увеличению задержек доступа и в итоге негативно отражается на производительности. В качестве некоторой компенсации увеличивают ассоциативность S-cache (обычно до 16 каналов), так как задержки при доступе к этому кэшу неважны. Например, согласно результатам исследований часто используемых целочисленных задач, у Intel Pentium III 16 Кбайт четырехканального D-cache было достаточно для покрытия около 93% запросов, а 16-Кбайт четырехканального I-cache — 99% запросов.

**Размер строки и тега кэш-памяти**

Немаловажная характеристика кэш-памяти — размер строки. Как правило, на одну строку полагается одна запись адреса (так называемый тег), которая указывает, какому адресу в оперативной памяти соответствует данная линия. Очевидно, что нумерация отдельных байтов нецелесообразна, поскольку в этом случае объем служебной информации в кэше в несколько раз превысит объем самих данных. Поэтому один тег обычно полагается на одну строку, размер которой обычно 32 или 64 байта (реально существующий максимум 1024 байта), и эквивалентен четырем (иногда восьми) разрядностям системной шины данных. Кроме того, каждая строка кэш-памяти сопровождается некоторой информацией для обеспечения отказоустойчивости: одним или несколькими битами контроля четности (parity) или восемью и более байтами обнаружения и коррекции ошибок (ЕСС, Error Checking and Correcting), хотя в массовых решениях часто не используют ни того, ни другого.

Размер тега кэш-памяти зависит от трех основных факторов: объема кэш-памяти, максимального кэшируемого объема оперативной памяти, а также ассоциативности кэш-памяти. Математически этот размер рассчитывается по формуле

Smg — размер одного тега кэш-памяти, в битах; Smem — максимальный кэшируемый объем оперативной памяти, в байтах; Scache — объем кэш-памяти, в байтах; А — ассоциативность кэш-памяти, в каналах.

Отсюда следует, что для системы с 1-Гбайт оперативной памятью и 1-Мбайт кэш-памятью с двухканальной ассоциативностью потребуется 11 бит для каждого тега. Примечательно, что собственно размер строки кэш-памяти никак не влияет на размер тега, но обратно пропорционально влияет на количество тегов. Следует понимать, что размер строки кэш-памяти не имеет смысла делать меньше разрядности системной шины данных, но многократное увеличение размера приведет к чрезмерному засорению кэш-памяти ненужной информацией и излишней нагрузке на системную шину и шину памяти. Кроме того, максимально кэшируемый объем кэш-памяти не обязан соответствовать максимально возможному устанавливаемому объему оперативной памяти в системе. Если возникнет ситуация, когда оперативной памяти окажется больше, чем может быть кэшировано, то в кэш-памяти будет присутствовать информация только из нижнего сегмента оперативной памяти. Именно такой была ситуация с платформой Socket7/Super7. Наборы микросхем для этой платформы позволяли использовать большие объемы оперативной памяти (от 256 Мбайт до 1 Гбайт), в то время как кэшируемый объем часто был ограничен первыми 64 Мбайт (речь идет о B-cache, находящемся на системной плате) по причине использования дешевых 8-бит микросхем теговой SRAM (2 бита из которых резервировалось под указатели действительности и измененности строки). Это приводило к ощутимому падению производительности.

Какая информация содержится в тегах кэш-памяти? Это информация об адресах, но как можно точно отобразить расположение строки кэш-памяти на всем пространстве кэшируемого объема оперативной памяти, используя столь незначительное количество адресных битов? Это понятие является фундаментальным в понимании принципов функционирования кэш-памяти.

Рассмотрим предыдущий пример, с 11-бит тегами. Учитывая логическое сегментирование благодаря двухканальной ассоциативности, можно рассматривать данную кэш-память как состоящую из двух независимых сегментов по 512 Кбайт каждый. Представим оперативную память как состоящую из «страниц» по 512 Кбайт каждая — их будет соответственно 2048 штук. Далее, log2 (2048) - 11 (основание логарифма равно 2, так как возможны только два логических состояния каждого бита). Это означает, что фактически тег — не номер отдельной строки кэш-памяти, а номер «страницы» памяти, на которую отображается та или иная строка. Другими словами, в пределах «страницы» сохраняется прямое соответствие ее «строк» с соответствующими строками кэш-памяти, т. е. п-я строка кэш-памяти соответствует п-й «строке» данной «страницы» оперативной памяти.

Рассмотрим механизм работы кэш-памяти разных видов ассоциативности. Допустим, имеется абстрактная модель с восемью строками кэш-памяти и 64 эквивалентными строками оперативной памяти. Требуется поместить в кэш строку 9 оперативной памяти (заметим, что все строки нумеруются от нуля и по возрастающей). В модели с прямым отображением эта строка может занять только одно место: 9 mod 8 = 1 (вычисление остатка от деления нацело), т. е. место строки 1. Если взять модель с двухканалыюй ассоциативностью, то эта строка может занять одно из двух мест: 9 mod 4=1, т. е. строку 1 любого канала (сегмента). Полноассоциативная модель предоставляет свободу для размещения, и данная строка может занять место любой из восьми имеющихся. Другими словами, фактически имеется 8 каналов, каждый из которых состоит из 1 строки.

Ни одна из вышеуказанных моделей не позволит, разумеется, поместить в кэш больше строк, чем он физически в состоянии разместить, они лишь предлагают различные варианты, различающиеся балансом эффективности использования кэша и скорости доступа к нему.

**Типы подключения кэш-памяти**

Количество портов чтения-записи кэш-памяти — показатель того, сколько одновременных операций чтения-записи может быть обработано. Хотя жестких требований и нет, определенное соответствие набору функциональных устройств ЦП должно прослеживаться, так как отсутствие свободного порта во время исполнения команды приведет к вынужденному простою.

Существует два основных способа подключения кэш-памяти к ЦП для чтения: сквозной и побочный (Look-Through и Look-Aside). Суть первого в том, что при необходимости данные сначала запрашиваются у контроллера кэш-памяти самого высокого уровня, который проверяет состояние подключенных тегов и возвращает либо нужную информацию, либо отрицательный ответ, и в этом случае запрос перенаправляется в более низкий уровень иерархии кэш-памяти или в оперативную память. При реализации второго способа чтения запрос одновременно направляется как кэш-контроллеру самого высокого уровня, так и остальным кэш-контроллерам и контроллеру оперативной памяти. Недостаток первого способа очевиден: при отсутствии информации в кэше высокого уровня приходится повторять запрос, и время простоя ЦП увеличивается. Недостаток второго подхода — высокая избыточность операций и, как следствие, «засорение» внутренних шин ЦП и системной шины ненужной информацией. Логично предположить, что если для кэшей L1 оптимальна сквозная схема, то для T-cache или B-cache побочная схема может оказаться более выгодной. Для S-cache выбор неоднозначен.

Различают также локальный и удаленный кэш. Локальным называют кэш, находящийся либо в ядре ЦП, либо на той же кремниевой подложке или в корпусе ЦП, удаленным — размещенный на системной плате. Соответственно локальным кэшем управляет контроллер в ядре ЦП, а удаленным — НМС системной платы. Локальный кэш с точки зрения быстродействия предпочтительнее, так как интерфейс к удаленному кэшу обычно мультиплексируется с системной шиной. С одной стороны, когда другой ЦП захватывает общую системную шину или какой-либо периферийный контроллер обращается к памяти напрямую, удаленный кэш может оказаться временно недоступным. С другой — такой кэш легче использовать в многопроцессорных системах.

Существуют два распространенных способа записи в кэш: сквозной (Write-Through) и обратной (Write-Back) записи. В первом случае информация одновременно сохраняется как в текущий, так и в более низкий уровень иерархии кэш-памяти (или прямо в оперативную память при отсутствии такового). Во втором — данные сохраняются только в текущем уровне кэш-памяти, при этом возникает ситуация, когда информация в кэше и оперативной памяти различается, причем последняя становится устаревшей. Для того чтобы при сбросе кэша информация не была необратимо потеряна, к каждой строке кэша добавляется «грязный» бит (dirty bit, иначе известный как modified). Он нужен для обозначения того, соответствует ли информация в кэше информации в оперативной памяти, и следует ли ее записать в память при сбросе кэша.

Также следует упомянуть способ резервирования записи (write allocation). При записи данных в оперативную память часто возникает ситуация, когда записываемые данные могут скоро понадобиться, и тогда их придется довольно долго подгружать. Резервирование записи позволяет частично решить эту проблему: данные записываются не в оперативную память, а в кэш. Строка кэша, вместо которой записываются данные, полностью выгружается в оперативную память. Так как вновь записанных данных обычно недостаточно для формирования полной строки кэша, из оперативной памяти запрашивается недостающая информация. Когда она получена, новая строка записывается, и тег обновляется. Определенных преимуществ или недостатков такой подход не имеет — иногда это может дать незначительный прирост производительности, но также и привести к засорению кэша ненужной информацией.

**TLB как разновидность кэш-памяти**

Почти все современные ЦП обладают TLB (Translation Look-aside Buffers, вспомогательные буфера преобразования). Своим существованием они обязаны тому факту, что ЦП в работе используют преимущественно виртуальные адреса оперативной памяти, в то время как контроллеры оперативной и кэш-памяти работают преимущественно с реальными адресами. Для того чтобы не вычислять при каждом обращении к памяти реальный адрес из виртуального, в ЦП присутствуют таблицы соответствия виртуальных адресов страниц памяти реальным. Как правило, их объем невелик (от единиц до сотен записей), но этого вполне достаточно, поскольку часто запрашиваемые данные или команды обычно хорошо локализуются в пределах страницы памяти размером 4 или 8 Кбайт.

Что же происходит, если запрашиваемого реального адреса какой-либо страницы не находится в TLB? A-box ЦП отрабатывает специальный вызов (exception trap), на который ОС должна адекватно отреагировать, т. е. произвести поиск нужной страницы в своих таблицах подсистемы виртуальной памяти. Если в процессе поиска окажется, что указанная страница находится в файле или разделе подкачки, то она должна незамедлительно быть оттуда считана в оперативную память. В итоге А-box ЦП получит реальный адрес нужной страницы памяти и процесс пойдет своим путем.

Стоит также заметить, что единой схемы адресации записей кэш-памяти не существует. Поэтому в зависимости от иерархического расположения и целевого назначения данной структуры кэш памяти, а также идей построения определенного ЦП может использоваться выборочно как реальная, так и виртуальная адресация записей, или даже гибридная схема (реальное тегирование с виртуальной индексацией), что предполагает и соответствующую организацию структуры TLB.

**Многопроцессорность**

**Причины появления многопроцессорных систем**

В настоящее время предельная производительность системы определяется мощностью ЦП и подсистемы памяти (контроллера памяти, который может быть встроен как в набор микросхем, так и непосредственно в процессор). Поэтому наиболее эффективное масштабирование по производительности обеспечивают именно многопроцессорные системы. Хотя кластерные решения, объединяющие несколько вычислительных узлов, как правило, при пересчете на один ЦП оказываются дешевле, но из-за накладных расходов на организацию их взаимодействия удельная производительность многопроцессорных машин все же заметно выше. Кроме того, использование многопроцессорных серверов позволяет минимизировать дублирование вспомогательных подсистем вычислительной машины — контроллеров ввода-вывода, дисковых массивов и т. п.

Нельзя не отметить и конструктивные преимущества многопроцессорных серверов: они обеспечивают максимальную удельную производительность при пересчете не только на один ЦП, но и на единицу объема, что в серверных платформах весьма важно.

Благодаря указанным преимуществам многопроцессорные платформы пользуются популярностью практически во всех сегментах серверного рынка. Двух- и четырехпроцессорные серверы младшего и среднего уровня имеют хорошее соотношение цена/производительность, и с точки зрения организации охлаждения они ненамного сложнее однопроцессорных. Решения с восемью и более процессорами имеют более сложную архитектуру, но и позволяют добиться максимальной производительности и надежности.

**Топологии многопроцессорных систем**

Типы многопроцессорных систем различаются прежде всего подходом к организации системной шины, связывающей, как и в однопроцессорных системах, ЦП и НМС. В настоящее время наиболее распространены три варианта архитектуры многопроцессорных систем:

общая шина (shared bus);

звездообразная топология (star topology), или топология «точка—точка» (point-to-point topology);

коммутируемая топология (switch-based topology). Наиболее простая — топология общей шины.

Как следует из названия, все ЦП в системе подключаются последовательно к общей системной шине и соответственно вынуждены делить между собой ее полосу пропускания. Этим определяются как преимущества, так и недостатки этой архитектуры. Одновременно с экономией на количестве проводников и общей стоимости разработки возникает проблема пропускной способности системной шины, который столь существенен, что при большом количестве ЦП (начиная с восьми) такая топология просто неприменима, так как задержки при обращении к памяти слишком велики из-за очереди на обращение к шине. Компенсацией недостаточной пропускной способности системной шины может служить добавление значительных локальных объемов кэш-памяти к каждому ЦП. Общая шина наиболее оправданна в системах с двумя процессорами. В четырехпроцессорных системах объемы кэш-памяти каждого ЦП приходится увеличивать в 2-4 раза, что приводит к существенному росту стоимости конечной системы.

В звездообразной топологии каждый ЦП располагает выделенной системной шиной, а поэтому отсутствует необходимость в разделении доступа. Системные шины всех ЦП подключаются непосредственно к набору микросхем. Хотя такая архитектура и позволяет минимизировать зависимость производительности ЦП от пропускной способности системной шины, но налагает очень высокие требования на НМС и системную плату в целом, усложняя их проектирование и увеличивая стоимость производства. Тем не менее и этот подход не позволяет подключить большое количество ЦП — звездообразная топология применяется, как правило, в системах с двумя или четырьмя процессорами.

Довольно интересный пример звездообразной топологии — четырехпроцессорные серверы Fujitsu PRIME-POWER 450. В их основу положен принцип скоростного универсального коммутатора, работающего на тактовой частоте 220 МГц, который и выполняет все основные функции набора микросхем. К нему по выделенным 128-бит 220-МГц шинам (110 МГц с применением технологии DDR) подключаются ЦП SPARC64 V с тактовой частотой 2 ГГц и выше. Далее к нему же подключаются два контроллера оперативной памяти (DDR SDRAM 110 МГц), каждый из которых располагает своим 128-бит каналом данных, и два моста шины PCI — по 110-МГц 64-бит каналам. Эта схема обеспечивает скоростной обмен данными между всеми узлами, но добавление, например, еще четырех ЦП и двух контроллеров оперативной памяти потребует серьезной перепроектировки коммутатора.

У систем с коммутируемой топологией нет проблем масштабируемости, свойственных вышеперечисленным шинным архитектурам. В этой архитектуре ЦП подключаются не к НМС (который может отсутствовать), а к коммутатору, соединенному с другими коммутаторами. Механизм подключения может быть как простым (широкая общая шина), так и иерархическим (древовидная иерархическая структура). На основании такой топологии можно проектировать системы с большим количеством ЦП (до 1024).

Рассмотрим коммутируемую топологию на примере 16-процессорных серверов HP AlphaServer GS160. В основе архитектуры лежит так называемый Quad Building Block (QBB), содержащий четыре ЦП Alpha 21264B (EV68C) с тактовой частотой до 1,22 ГГц, до 16 Мбайт B-cache на каждый ЦП, четыре модуля регистровой SDRAM ЕСС общим объемом до 32 Гбайт, два модуля стабилизатора питания, два 64-бит 33-МГц РСI-интерфейса, собственно коммутатор и интерфейс GlobalPort для связи с другими коммутаторами.

Четыре таких QBB объединяют в двухуровневую древовидную структуру через главный коммутатор, находящийся на отдельном модуле и обеспечивающий пропускную способность между двумя отдельно взятыми QBB по 1,75 Гбайт/с в каждом направлении.

Особняком стоят системы, в которых контроллер оперативной памяти, а часто и контроллеры периферийных шин интегрированы в ядро ЦП, а поэтому необходимость в системной шине отпадает. Такие системы строятся по NUMA-архитектуре, о которой пойдет речь дальше. Их можно рассматривать как представителей коммутируемой топологии.

**Доступ к памяти в многопроцессорных системах**

Независимо от количества ЦП в системе и топологии системной шины, им придется разделять доступ к общему пространству оперативной памяти, за исключением случаев, когда используется сегментирование аппаратных ресурсов некоторого сервера или суперкомпьютера для запуска нескольких копий ОС. В настоящее время существуют три типа доступа к оперативной памяти в многопроцессорных системах:

классический SMP (Symmetrical Multi-Processing, симметричная многопроцессорность);

распределенный SMP (distributed SMP);

NUMA (Non-Uniform Memory Access, неоднородный доступ к памяти).

Первоначальная модель симметричней многопроцессорности, теперь называемая классической, предусматривала, что все ЦП в системе имеют доступ к физически единому пространству оперативной памяти, обслуживаемому специальным контроллером. Все ЦП равнозначны по уровню прав доступа, какие-либо механизмы приоритетности отсутствуют, системная логика обрабатывает запросы в порядке поступления, и только на нее возложены все функции по работе с оперативной памятью. Соответственно задержки при доступе к памяти одинаковы для всех ЦП при обращении к любой области памяти. Это наиболее простая модель организации доступа к памяти в многопроцессорных системах, но не самая производительная и не позволяющая подключать большое количество ЦП из-за объективных проектировочных ограничений.

Принцип распределенной многопроцессорности предполагает формирование из имеющихся в системе ЦП логических групп, каждая из которых располагает своим (локальным) объемом оперативной памяти. Естественно, что для обеспечения целостности и нормального функционирования всего пространства оперативной памяти эти логические группы должны быть как-то связаны между собой. Задержки при доступе к памяти будут соответственно различны в зависимости от того, происходит обращение к локальной или удаленной области памяти.

Следует понимать, что, несмотря на некоторое сходство, распределенная SMP — не частный случай коммутируемой архитектуры: в последней могут быть реализованы как SMP, так и NUMA. Например, в вычислительном комплексе Intel Profusion восемь ЦП Intel Xeon разделены на две группы по четыре ЦП и связаны коммутатором, к которому подключается НМС с контроллером оперативной памяти. Таким образом, при обращении к памяти все процессоры равноправны и Intel Profusion представляет собой классическую SMP-систему. А вот описанный выше сервер HP AlphaServer GS160 спроектирован по принципу распределенной многопроцессорности.

Принцип неоднородного доступа (NUMA) состоит в явном разделении на локальную и удаленную память для каждого ЦП. Так как главный недостаток коммутируемой архитектуры — относительно низкая скорость обмена информацией между коммутаторами, возникла идея интегрировать контроллер памяти в ядро ЦП. Модули оперативной памяти в этом случае подключаются напрямую к ЦП (без НМС). В многопроцессорной системе такому ЦП придется запрашивать информацию из оперативной памяти других ЦП и отвечать на их запросы. Модель памяти называется неоднородной, потому что скорость доставки данных из локальной оперативной памяти ЦП будет значительно выше,, чем из оперативной памяти, подключенной к другому ЦП, причем разница будет все более ощутимой с увеличением количества ЦП в системе.

Модели с распределенной многопроцессорностью и неоднородным доступом позволяют создавать более масштабируемые решения, чем классическая SMP. Однако они предъявляют некоторые специфические требования к ОС, которая должна оптимизировать загрузку процессоров таким образом, чтобы процесс, выполняемый на одном ЦП, как можно реже использовал информацию, находящуюся в локальной оперативной памяти другого ЦП.

**Современные серверные процессоры**

**Intel Xeon**

Сегодня наиболее распространенные серверные ЦП — семейство Intel Xeon с архитектурой NetBurst, выполненные по технологии Hyper-Threading (все современные), совместимые по набору команд с другими ЦП архитектуры х86. Фактически это серверные версии соответствующих ЦП семейства Pentium 4, отличающиеся возможностью работы в многопроцессорном режиме и корпусом Socket 604 (ранее Socket 603). В семейство входят два типа процессоров: Xeon DP (Dual Processor) и Xeon MP (Multi Processor). Как следует из названий, Xeon DP может быть использован максимум в двухпроцессорных конфигурациях, а Xeon МР — в четырехпроцессорных. Впрочем, с помощью технологий, подобных Intel Profusion, указанные ограничения могут быть преодолены. Все ЦП рассчитаны на системы с топологией общей шины. В настоящее время существенных отличий между Xeon DP и Xeon MP не наблюдается, различаются только частоты системной шины: 200 и 166 МГц соответственно.

Архитектура NetBurst обеспечивает очень высокие тактовые частоты благодаря длинным конвейерам основных ФУ — 20 стадий целочисленных конвейеров для ЦП, изготавливаемых по 180-нм и 130-нм проектным нормам и 31 стадия для ЦП, выполненных по 90-нм техпроцессу (в обоих случаях не считая стадий декодирования). Так как все ЦП архитектуры NetBurst (Xeon, Pentium 4 и Celeron) имеют общую функциональную логику, незначительно изменяемую при переходе на новые проектные нормы, то все те недостатки, присущие ядрам ЦП Pentium 4 (Willamette, Northwood, Prescott), свойственны и модельному ряду Xeon. Известно, что для достижения стабильной работы на высоких частотах конвейерам NetBurst часто приходится отрабатывать «холостые» такты для синхронизации соседних стадий. Кроме того, большая длина конвейера делает его более чувствительным к ошибкам при предсказании условных переходов, так как больше времени уходит на сброс и перезаполнение конвейера.

Декодер х86-команд не синхронизирован с конвейерами ФУ и производит выборку (в том числе и упреждающую) в кэш микрокоманд (trace cache) со скоростью одна команда х86 за такт. Это оказывается эффективным при многократном исполнении одного и того же кода с достаточной временной локализацией (нет необходимости декодировать команды многократно), хотя упреждающее декодирование может засорять кэш. Согласно данным Intel, размер этого кэша 12 288 микрокоманд. Данные о длине микрокоманд в документации не представлены, но, по некоторым данным, она составляет до 10 байт в зависимости от версии ядра. Если же учесть, что после декодирования одна команда х86 преобразовывается обычно в две-три внутренние микрокоманды, то кэша микрокоманд хватит для хранения 4—6 тыс. х86-команд.

Ядро NetBurst имеет Е-box, состоящий из трех конвейеров (два для выполнения простых микрокоманд и один для обработки сложных и редко используемых). При этом два «простых» конвейера входят в состав так называемого Rapid Execution Engine, работающего на удвоенной частоте ядра.

А-box также состоит из трех конвейеров (один обрабатывает микрокоманды загрузки целочисленных данных, один — сохранения целочисленных данных и один — загрузки/сохранения/перемещения вещественных данных). Оба адресных конвейера по работе над целочисленными данными также являются частью Rapid Execution Engine. F-box представлен одним универсальным конвейером, предназначенным для проведения расчетных операций над командами из наборов х87, ММХ и SSE. Обслуживают перечисленные конвейеры шесть планировщиков, осуществляющих выборку из кэша микрокоманд и помещающих их в соответствующие очереди. Каждый планировщик обслуживает отдельный конвейер: по одному на каждый «быстрый» целочисленный, один — на комплексный целочисленный, один — на оба целочисленных адресных, один — на вещественный адресный и один — на вещественный конвейер. Вполне естественно, что оба планировщика, обслуживающих целочисленные конвейеры, — часть Rapid Execution Engine, в отличие от обслуживающего целочисленные адресные конвейеры — он работает на частоте ядра ЦП. Для увеличения производительности целочисленных конвейеров файл их регистров работает на удвоенной частоте ядра ЦП.

Благодаря тому что оба «простых» конвейера Е-bох работают на удвоенной частоте ядра ЦП, многие простые операции (сложение, вычитание, логические операторы) теоретически могут быть выполнены за половину такта. Однако следует учитывать, что эти конвейеры работают с 16-бит операндами, поэтому на 32-бит операции тратится полный такт (сначала обрабатываются младшие 16 бит, а затем старшие 16). Впрочем, предусмотрен режим back-to-back, при котором одновременно на одном конвейере будет рассчитываться младшая часть результата, а на другом — старшая. Таким образом, теоретическая возможность выполнения полной операции в пределах полутакта сохраняется.

В то же время некоторые операции, например, смещения (shifts), выполняются за четыре такта, так как проходят по относительно медленному «сложному» конвейеру. Необходимо отметить, что в ЦП, изготавливаемых по 130-нм проектным нормам, смещения обрабатываются на «простых» конвейерах.

Перечисленные особенности привели к тому, что первые ЦП- семейства Intel Xeon нередко уступали по производительности процессорам семейства Pentium III, хотя и работали на более высокой тактовой частоте. Частично решить эту проблему позволила перекомпиляция программной базы с оптимизацией под архитектуру Net-Burst. В настоящее время, разумеется, сравнивать ЦП архитектуры Р6 и NetBurst неправомерно — частота и производительность процессоров семейства Xeon выросли до высочайшего уровня, и, кроме того, эти ЦП имеют гораздо большие функциональные возможности.

Одно из преимуществ архитектуры NetBurst — широкая (256 бит, не включая биты ЕСС) шина данных между D-cache и S-cache (для которого Intel использует название Advanced Transfer Cache). Поскольку D-cache мал и работает в режиме сквозной записи, то именно S-cache является основным кэшем архитектуры NetBurst. Особенно важен низкий уровень задержек при обращении к нему: девять тактов, в том числе два на доступ к D-cache, так как S-cache подключен по схеме look-through. Впоследствии, при проектировании ядер для 90-нм проектных норм, задержки доступа значительно увеличились — 15 тактов, в том числе четыре на доступ к D-cache. Это отразилось на производительности, причем довольно часто среди ЦП с равной частотой, изготовленных по 130-и 90-нм технологическому процессу, мощнее оказывалась именно первая версия. Стоит отметить, что при доступе к D-cache со стороны вещественного конвейера А-Ьох задержки еще более высоки: девять тактов для 180- и 130-нм ЦП и 12 тактов для 90-нм ЦП.

Таким образом, S-cache оказывает огромное влияние на производительность ЦП Xeon. Он является включающим (inclusive), т. е. содержит полную копию содержимого D-cache. К тому же именно в S-cache хранится тело исполняемого процесса и, если учесть высокие тактовые частоты, на достижение которых архитектура была изначально ориентирована, становится очевидно, что увеличение размера S-cache неминуемо приведет к росту производительности на большинстве задач. Именно поэтому при переходе на новые проектные нормы его размер неизменно увеличивался (с 256 Кбайт у Foster до 2 Мбайт у Irwindale).

Кроме того, в некоторых моделях Xeon используется T-cache, также интегрированный в ядро, но гораздо более медленный (асинхронный) по сравнению с S-cache. Он тоже включающий относительно S-cache. Его основная задача — снизить частоту обращений к оперативной памяти, что очень важно, учитывая топологию общей шины.

Еще одной интересной особенностью архитектуры стала технология Hyper-Threading, позволяющая организовать два логических ЦП на основе одного физического. Своим существованием эта технология обязана тому, что в современных ЦП при исполнении кода какого-либо приложения многие ФУ простаивают, и их можно использовать для обработки второго потока. Для успешной реализации такого подхода требуется некоторая доработка декодера команд и наличие дополнительного набора основных регистров, а также совместимость программных средств — ОС и компилятора. Для работы второму логическому ЦП выделяется половина физических кэшей: D-cache, S-cache и T-cache (если имеется), вследствие чего включение Hyper-Threading при использовании приложений, не совместимых с. этой технологией, приводит к некоторому падению производительности. При использовании перекомпилированных (с оптимизацией под Hyper-Threading) приложений наблюдается некоторый рост производительности, величина которого зависит от специфики решаемых задач.

С переходом на 90-нм проектные нормы существенного роста тактовой частоты ЦП Intel Xcon не произошло — рубеж в 4 ГГц все еще не взят. Хотя более высокочастотные версии ЦП на том же техпроцессе и были заявлены, но их выпуск может и не состояться, так как недавно был анонсирован двухъядерный Pentium 4 Extreme Edition 840 (Smithfield) с тактовой частотой 3,2 ГГц. Ожидается, что на его ядре будут выпущены процессоры Xeon DP (Dempsey) и Xeon МР (Paxville).

**AMD Opteron**

Процессоры семейства Opteron стали заметным шагом компании AMD на серверном рынке. Их предшественниками были ЦП семейства Athlon MP, которые хотя и пользовались определенной популярностью (благодаря использованию архитектуры «точка—точка», предоставлявшей каждому ЦП полноценную системную шину), не могли быть использованы для создания систем более чем с двумя процессорами.

ЦП семейства Opteron построены на ядре К8 (AMD64), как и процессоры Athlon 64, и имеют некоторые особенности, необходимые для работы в многопроцессорных системах. ЦП Opteron оснащаются встроенным двухканальным контроллером регистровой DDR SDRAM со 128-бит шиной памяти (не включая 16 бит для ЕСС), этот контроллер позволяет применять модули памяти с частотой до 200 МГц.

Opteron совместимы с 64-бит расширениями, что, в отличие от рынка ПК, для серверных платформ очень важно. Уже довольно продолжительное время некоторые черты 32-бит архитектуры х86 препятствовали проникновению процессоров х86 в сегмент серверов среднего уровня. Причина проста: недостаток адресного пространства — 32-бит процессор может адресовать лишь 4 Гбайт ОЗУ. Такой объем адресного пространства в настоящее время не может считаться достаточным даже для некоторых серверов младшего уровня.

С выпуском ЦП Pentium Pro в 1995 г. компания Intel попробовала решить эту проблему с помощью технологии РАЕ36 (Page Address Extension, расширение страничной адресации), позволявшей расширить адресное пространство до 64 Гбайт путем использования дополнительного 4-бит сегментного регистра. Но работа с оперативной памятью в таком режиме значительно снижала производительность (в некоторых ситуациях почти вдвое). 64-бит адресация решает эту проблему без какого-либо ущерба для производительности.

Теоретически Opteron-системы относятся к архитектуре NUMA: в каждом процессоре для связи с соседними ЦП и периферией предусмотрено от одной до трех высокоскоростных двунаправленных шин HyperTrans-port. Однако на практике разница в задержках при обращении к локальной и удаленной памяти невелика — всего примерно 30% (в то время как в классических NUMA-системах она достигает сотен процентов), что позволяет компании AMD говорить о SUMO-архитектуре (существенно однородная архитектура памяти).

AMD делит ЦП семейства Opteron на три группы:

Opteron lxx, для однопроцессорных серверов;

Opteron 2xx, для двухпроцессорных серверов;

Opteron 8xx, для четырех- и восьмипроцессорных серверов.

Для связи с другими ЦП в многопроцессорной конфигурации и набором микросхем используются скоростные шины HyperTransport — процессоры Opteron оснащены тремя такими шинами. Шина HyperTransport — это двунаправленная последовательная шина, состоящая из 16 каналов и работающая на частоте 800 МГц. Пропускная способность шины в синхронном режиме составляет 3,2 Гбайт/с благодаря использованию технологии DDR. Шина HyperTransport может работать в синхронном (когерентном) и асинхронном режимах. Первый предназначен для связи с другими ЦП, второй — для связи с периферийным контроллером. Фактически ядра ЦП Opteron всех серий идентичны, но в ЦП серии 8хх активированы все три шины HyperTransport, в серии 2хх — только две, а в серии lxx — одна.

По сравнению с предыдущим поколением ЦП Athlon МР дизайн ядра Opteron конструктивно изменился. Кроме уже упомянутого встроенного контроллера оперативной памяти, шин HyperTransport и средств AMD64, были реализованы и менее масштабные доработки. Например, длины целочисленных и вещественных конвейеров увеличены на две стадии, с 10 до 12ис 15 до 17 соответственно. Это было вызвано в основном необходимостью переработки механизма декодирования команд, хотя количество декодеров осталось неизменным и равным трем. При декодировании наиболее простые и часто используемые CISC-команды х86 преобразуются аппаратной логикой ЦП в одну или две внутренние микрокоманды, которые затем группируются и направляются на ФУ. Сложные и нечасто используемые команды преобразуются программной логикой ЦП, основанной на загружаемом при инициализации ЦП микрокоде, что требует существенно больше времени и большего количества внутренних микрокоманд в расчете на одну декодируемую. Примечательно, что именно благодаря наличию программируемой части удалось легко дополнить ЦП Athlon XP и Athlon MP с ядром Palomino средствами SSE.

Кроме того, большее количество команд декодируется аппаратной логикой. Это относится прежде всего к операциям с вещественными числами — наборам команд SSE и SSE2. Число целочисленных команд, обрабатываемых программной логикой, сократилось на 8%, а вещественных — на 28%. Количество конвейеров ФУ осталось неизменным: по три на Е-Ьох, А-box и F-box. Последний, как и прежде, обрабатывав SIMD-команды, поступающие в виде наборов команд ММХ, Extended ММХ, 3DNow!, Extended 3DNow!, SSE и SSE2 (процессоры, изготовленные по 90-нм проектным нормам, начиная с ядра версии Е, совместимы и с SSE3). Необходимо отметить, что поскольку внутренняя разрядность вещественных конвейеров не изменилась (она равна 64 бит), то оптимизированный под SSE, SSE2 или SSE3 код не дает практически никакого выигрыша в производительности по сравнению с 3DNow! или Extended 3DNow!. Одна команда, оперирующая 128-бит векторизированными данными (состоящими из четырех слов одинарной точности или двух двойной), будет неминуемо разложена на две команды, оперирующие 64-бит векторизованными (два слова двойной точности) или скалярными данными (одно слово двойной точности). Они будут обработаны в порядке очереди, т. е. в два прохода.

Следует отметить и модернизированный блок С-bох: основные характеристики I-cache, D-cache и S-cache не изменились, и размер линии остался равным 64 байт, но узкая 64-бит шина к S-cache была заменена на две однонаправленные 64-бит шины. Доступ к S-cache (в том числе три такта для доступа к D-cache) занимает 11 тактов, а максимальная задержка составляет 16 тактов (в ситуации, когда строку D-cache, на место которой записывается новая строка из S-cache, необходимо записать в S-cache, но victim buffer, служащий в подобных ситуациях временным хранилищем, переполнен). Емкость таблицы истории переходов возросла с 4 до 16 Кбайт.

Благодаря высокой производительности, удачной архитектуре, хорошей масштабируемости, совместимости с 64-бит расширениями и отсутствию падения производительности при исполнении 32-бит кода ЦП семейства Opteron были благоприятно восприняты на рынке. Серверы на основе этих ЦП успешно выпускаются IBM, Sun, Fujitsu Siemens, HP и другими компаниями. Вслед за дебютом двухъядерного Athlon 64 Х2 (с ядром Toledo, Manchester) ожидается выход серверов на основе двухъядерных Opteron x65, х70 и х75 (ядра Egypt, Italy и Denmark).

**Intel Itanium**

Совсем недавно процессоры семейства Intel Itanium позиционировали в качестве единственно возможного преемника устаревающей архитектуры х86, как архитектуру post-RISC, способную оставить позади все существующие традиционные RISC-архитектуры, о которых пойдет речь дальше. Однако архитектура х86 до сих пор не сходит со сцены и пользуется огромной популярностью. Не способствует массовому распространению ЦП семейства Itanium соотношение цена/производительность. Тем не менее эти процессоры пользуются определенным успехом и востребованы на рынке серверов среднего уровня.

История семейства Itanium уходит корнями в 1992-93 гг., когда после отказа Роберту Палмеру (Robert Palmer), главному исполнительному менеджеру компании DEC, в сотрудничестве по продвижению и разработке 64-бит RISC-процессоров Alpha Эндрю Гроув (Andrew Grove), главный исполнительный менеджер компании Intel, инициировал создание стратегического альянса с Hewlett-Packard для совместной разработки и продвижения собственной 64-бит RISC-архитектуры. Одновременно Intel продолжила развивать модельный ряд ЦП архитектуры х86, выпустив процессоры 486 и Pentium. Со своей стороны, Hewlett-Packard работала над фирменными процессорами PA-RISC.

Первый процессор Itanium (кодовое название Merced) появился лишь в 1999 г. — компания Intel начала поставки ЦП своим крупнейшим партнерам. Столь длительная задержка выпуска (первая информация о ядре Merced появилась еще в 1997 г.) была связана в основном с технологическими причинами: нестабильностью работы на проектной тактовой частоте 800 МГц. Официальный анонс нового семейства состоялся лишь в июле 2001 г.

Системы предоставлялись для ознакомительных целей и отладки программного обеспечения, чтобы обеспечить более благоприятную встречу следующему представителю семейства — процессору с ядром McKinley, официально представленному под именем Itanium 2 в июле 2002 г. Позже появились обновленные версии ядра — Madison в июле 2003 г. и Deerfield в сентябре того же года.

64-бит процессоры семейства Itanium выполнены по архитектуре VLIW, которая ориентирована на повышение количества выполняемых команд в расчете на один такт ЦП. После декодирования команды, упакованные в 16-байт пакеты по три в каждом, направляются на ФУ процессора. Декодеры в состоянии обрабатывать до шести команд за один такт ЦП. Планировщик ЦП сравнительно примитивен: вся работа по созданию оптимального потока команд возлагается на программные средства, т. е. на компиляторы и оптимизаторы, хотя иногда планировщик может несколько скорректировать поток команд.

Количество имеющихся конвейеров впечатляет: четыре в Е-box, два двойной и два одинарной точности в F-box, два в А-box, четыре в блоке SIMD-команд и три в блоке обработки условных переходов. Соответственно используется очень большое число регистров: 128 целочисленных, 128 вещественных и 72 специальных временных. Суть применения последних состоит в том, что при обработке условного перехода логика ЦП рассчитывает сразу обе ветви и хранит результаты во временных регистрах до тех пор, пока не будет достоверно известно, какая из ветвей окажется верной, тогда нужные результаты будут сразу же перенесены в основные регистры. После декодирования пакеты команд попадают в так называемый «буфер разъединения» (decoupling buffer), емкость которого составляет восемь пакетов, оттуда каждая отдельная команда направляется в один из девяти доступных конвейерных портов. Длина конвейеров — 10 стадий у Merced, а у McKinley и последователей она сокращена до восьми. Применяется техника переименования регистров, дополненная «складированием» и «ротацией» регистров.

Все ЦП семейства Itanium имеют три уровня кэш-памяти. Merced располагает 16-Кбайт I-cache с четырехканальной ассоциативностью, 16-Кбайт двухпортовым D-cache с четырехканальной ассоциативностью и сквозной записью, 96-Кбайт S-cachc с шестиканалыюй ассоциативностью и обратной записью. Также подключается внешний B-cache с четырехканальной ассоциативностью и обратной записью объемом 2 или 4 Мбайт. I-cache полностью конвейеризирован и в состоянии поставлять 32 байт команд (два пакета) за один такт ЦП, он дополнен полноассоциативным TLB на 64 записи. Задержки при чтении (load-to-use latency) из I-cache и D-cache всего два такта. Размер строки I-cache и D-cache — 32 байт, a S-cache и B-cache — 64 байт. D-cache располагает TLB на 32 записи, a S-cache — на 96 записей. Ширина шины данных к S-cache 256 бит. B-cache подключается к ядру ЦП через 128-бит шину данных, работающую на частоте ядра ЦП. Частота системной шины 133 МГц (с использованием технологии DDR — 266 МГц), а ширина канала данных 64 бит (с дополнительными битами для ЕСС или parity). Теоретическая максимальная скорость обмена данными по шине 2,1 Гбайт/с.

Ядро McKinley — это эволюционное развитие ядра Merced. Количество конвейеров А-box удвоено (четыре). Размер строки I-cache и D-cache увеличен до 64 байт, кроме того, D-cache стал четырехпортовым (по два порта для чтения и записи). Размер S-cache увеличен до 256 Кбайт, уровень ассоциативности вырос до восьми, а размер строки достиг 128 байт. Внешний B-cache был упразднен, а вместо него в ядро ЦП встроен 3-Мбайт T-cache с 12-канальной ассоциативностью и обратной записью, существуют также модели с 1,5-Мбайт T-cache. Ширина канала данных к T-cache расширена до 256 бит (с дополнительными 32 бит для ЕСС). Размер строки T-cache составляет 128 байт. Время доступа к S-cache пять и более тактов, к T-cache — 12 и более тактов. Количество записей в TLB S-cache увеличено до 128. Частота системной шины возросла до 200 МГц (эффективная частота — 400 МГц), а ширина канала данных до 128 бит (с дополнительными битами для ЕСС или четности), что дало возможность добиться теоретической максимальной скорости обмена данными по шине 6,4 Гбайт/с.

В ядро Madison было внесено меньше изменений: оно изготовлено по 130-нм технологии, оснащается 6-Мбайт T-cache, кроме того, существует модификация Madison 9М, у которой размер T-cache увеличен до 9 Мбайт. Еще одна модификация Madison — ядро Deerfield — обладает только 3-Мбайт T-cache и предназначено для двухпроцессорных конфигураций.

ЦП семейства Itanium подключаются к общей шине (это традиционно для процессоров Intel), максимальное количество ЦП на одной шине ограничено четырьмя. На практике пропускной способности системной шины в четырехпроцессорных конфигурациях часто оказывается недостаточно, поэтому при проектировании серверов с большим количеством ЦП часто ставят только по два ЦП на одну шину и соединяют отдельные шины при помощи скоростных коммутаторов.

Все ЦП семейства Itanium совместимы с набором команд х86 на аппаратном уровне. Однако производительность при выполнении х86-кода очень невысока — она сравнима с мощностью первых моделей Pentium. Это объясняется тем, что концептуальные различия между архитектурами настолько значительны, что логика ЦП самостоятельно не в состоянии обеспечить должную нагрузку на ФУ, чтобы вывести производительность на приемлемый уровень. Позже компания Intel выпустила связующее ПО (middleware) для Windows, которое на лету транслировало х86-команды в код Itanium с последующей оптимизацией. Этот подход позволил добиться падения производительности только в два раза по сравнению с изначально откомпилированным для архитектуры Itanium кодом. Не исключено, что вскоре Intel будет вынуждена доработать это ПО для обеспечения совместимости с 64-бит расширениями х86.

При сравнении архитектуры Itanium с NetBurst видны разные подходы к достижению максимальной производительности: если в NetBurst ставка делается на высокие тактовые частоты ядра ЦП при относительно небольшом количестве конвейеров ФУ и их большой длине, то в Itanium имеет место обратный вариант.

В настоящее время Itanium 2 лидирует по производительности на вещественных операциях и показывает довольно хорошую производительность на целочисленных. Ожидается выход двухъядерного Montecito с Hyper-Threading и 24-Мбайт встроенным T-cache — этот ЦП поставит рекорд по количеству транзисторов в ядре — около 1,7 млрд.

**DEC Alpha**

История этой легендарной архитектуры, на протяжении 10 лет показывавшей наивысшую производительность как на целочисленных, так и на вещественных операциях, началась с 1982 г., когда в Западной исследовательской лаборатории Digital Equipment Corporation (DF.C) началась работа над RISC-проектом Titan. Несколько позже были начаты еще три независимых RISC-проекта в рамках компании, которые в 1985 г. были объединены в проект PRISM (Parallel Instruction Set Machine, машина с параллельным набором команд). В 1988 г. проект был неожиданно закрыт, хотя уже через год началась работа над другой архитектурой ЦП, которая унаследовала многие особенности PRISM. Эта 64-бит архитектура получила название Alpha, а первый ЦП на ее основе, представленный в ноябре 1992 г., назывался Alpha 21064 и имел кодовое наименование EV4. Стартовав на очень высоких по меркам того времени тактовых частотах (150-200 МГц), сбалансированная архитектура этого ЦП привлекла внимание. Однако многочисленные просчеты руководства компании в вопросах ценообразования и позиционирования на рынке привели к тому, что архитектура так и не получила широкого распространения. Ситуацию не исправили ни выпуск недорогой модели 21066 (LCA4, первый ЦП со встроенным контроллером оперативной памяти), ни переход на новый четырехслойный 0,5-мкм техпроцесс и рост тактовых частот: в 1993 г. были представлены модели 21064А (EV45) с тактовой частотой 300 МГц и 21066А (LCA45) с тактовой частотой 233 МГц. В 1994 г. появился ЦП следующего поколения — 21164 (EV5) с удвоенным количеством конвейеров ФУ и встроенным 96-Кбайт S-cachc. Так как проектные нормы остались прежними, тактовая частота EV5 увеличилась лишь до 333 МГц. EV5 также оказался довольно дорогим в производстве из-за большого количества транзисторов (9,3 млн.) и площади ядра (299 мм2). После перепроектировки доработанного EV5 под 0,35-мкм технологию в 1995 г. появилась модификация 21164А (EV56), достигшая тактовой частоты 667 МГц. В 1997 г. был выпущен упрощенный вариант (без S-cacho.) этого ЦП — 21164РС (РСА56), он позиционировался как массовый, но производительный ЦП.

В январе 1998 г. DEC, испытывавшая серьезные финансовые трудности, была куплена компанией Compaq, одним из ведущих сборщиков компьютерной техники с архитектурой х86. Сформировав альянс с Samsung, эта компания пыталась продвигать на рынке последнюю разработку DEC — процессор 21264 (EV6), который хотя и изготавливался по тем же проектным нормам, что и EV56, показывал почти вдвое более высокую производительность при равных тактовых частотах. Это оказалось возможным благодаря использованию нового ядра, в котором было реализовано внеочередное выполнение команд, новая схема подключения кэш-памяти и новая системная шина DDR. Именно эта системная шипа была позже лицензирована AMD и применялась в процессорах семейства Athlon.

Однако даже столь удачная архитектура и внедрение новых проектных норм, позволивших значительно поднять тактовые частоты и снизить себестоимость, были выпущены 21264А (250-нм, EV67) и 21264В (18-нм, EV68A и EV68C) с частотами до 1250 МГц — не помогли этим ЦП завоевать популярность. Compaq и Sam sung не сумели обеспечить эффективный сбыт, ограничив область применения серверами и дорогими рабочими станциями. В июне 2001 г. Compaq объявила о переводе всех своих серверных решений на архитектуру Itanium, одновременно остановив разработку 21464 (EV8). В январе 2002 г. уже под маркой HP (поглотившей в конце 2001 г. Compaq) был анонсирован последний процессор Alpha — 21364 (EV7). Поскольку компания HP в сотрудничестве с Intel работала над развитием Itanium-платформ и располагала собственной архитектурой РА-RISC, архитектура Alpha оказалась «третьим лишним». Формально она была «похоронена» в августе 2004 г., когда компания Hewlett-Packard объявила о своих стратегических планах на рынке серверов, однако ее развитие прекратилось задолго до этого. Продукция на основе архитектуры Alpha перестанет продаваться под торговой маркой HP в 2006 г.

Тем не менее представители двух последних поколений архитектуры Alpha доступны на рынке и сейчас. В основе ядра EV6 лежит четырехконвейерный Е-box и двухконвейерный F-box (длина всех шести конвейеров 7 стадий). Функции А-box по подсчету виртуальных адресов возложены на Е-box. Кроме стандартных целочисленных команд и BWX (Byte-Word Extension), E-box поддерживает набор векторных команд MVI (Motion Video Instructions), предназначенных для ускорения расчетов, связанных с обработкой видео- и аудиопотоков. Процессор имеет 80 физических целочисленных и 72 физических вещественных регистра, однако количество логических (т. е. адресуемых программно) регистров осталось тем же, что и в предыдущих поколениях архитектуры: 32 целочисленных и 32 вещественных. 64-Кбайт I-cache с двухканальной ассоциативностью позволяет считывать по четыре команды за такт. Двухпортовый D-cache имеет объем 64 Кбайт, деухканальную ассоциативность и обратную запись. Также предусмотрен B-cache с прямым отображением и обратной записью, объем которого до 16 Мбайт. Он состоит из высокоскоростных микросхем статической памяти (обычно с отложенной записью) и подключается к ядру по выделенной шине со 128-бит каналом данных. Частота работы B-cache программируется и может находиться в диапазоне от 1/8 до 2/3 частоты ядра ЦП. Как уже упоминалось, системная шина использует технологию DDR для передачи данных, ее разрядность 64 бита (с дополнительными 8 битами для ЕСС).

EV7 представляет собой модификацию EV6, в которой оставшееся без изменений ядро дополнено 1,75-Мбайт S-cache с семиканальной ассоциативностью и обратной записью (подключаемым по той же шине со 128-бит каналом данных), двумя пятиканальными контроллерами оперативной памяти стандарта Direct Rambus (четыре 16-бит канала данных и один 16-бит канал для избыточной отказоустойчивости), четырьмя контроллерами для связи с соседними ЦП (с пропускной способностью свыше 6 Гбайт/с на канал) и одним контроллером для связи с периферией. Все вышеперечисленные контроллеры EV7 работали на 2/3 частоты ядра ЦП.

EV6 и его модификации использовали звездообразную топологию системной шины, в то время как системы на процессорах EV7 и EV7z строились по NUMA-топологии, маршрутизация данных между ЦП осуществляется встроенными коммутаторами. На практике к одному набору микросхем (DEC Tsunami, он же Typhoon) подключалось не более четырех ЦП семейства EV6, поэтому для построения систем с большим количеством ЦП использовались коммутируемые решения (например, уже рассмотренный AlphaServer GS160). На базе ЦП семейства EV7 могут быть созданы многопроцессорные серверы с количеством процессоров до 128. Благодаря четырем встроенным контроллерам каждый ЦП этого семейства может быть подключен к четырем соседним ЦП по произвольным алгоритмам.

С момента выхода EV6 в 1998 г. никаких изменений в функциональную часть ядра ЦП не вносилось, кроме того, для создания серверов на этом ЦП применяется НМС Tsunami, разработанный еще инженерами DEC. Его характеристики не удовлетворяют современным требованиям — частота системной шины составляет 83 МГц, на этой же частоте работает и контроллер памяти SDRAM (ширина шины памяти 256 или 512 бит). Частично недостатки НМС компенсирует большой объем B-cache, но по уровню производительности серверы на базе EV6—EV7 уступают современным системам.

**Hewlett-Packard PA-RISC**

Архитектура PA-RISC, известная как HP/PA (PA, Precision Architecture, точная архитектура), была разработана компанией Hewlett-Packard в конце 80-х гг. прошлого века. Руководство компании приняло решение перевести выпускаемую компьютерную технику с CISC-архитектур, таких, как 32-бит Motorola 68K и 16-бит собственной разработки, на потенциально перспективную свою же RISC-архитектуру.

Прообразом архитектуры РА-RISC считается 32-бит архитектура FOCUS, впервые воплощенная в рабочей станции HP 9000/520, которую в августовском выпуске Hewlett-Packard Journal 1983 г. назвали «настольным мэйнфреймом для инженеров и ученых». Каждый из максимум трех 18-МГц ЦП этой рабочей станции состоял из пяти микросхем, изготавливаемых по технологическому процессу NMOS3. Примечательно, что эта рабочая станция комплектовалась HP-UX, многопользовательской UNIX-подобной ОС, которая дожила до наших дней.

Первая спецификация на новую архитектуру, PA-RISC 1.0, заложила ее основы:

длина любой команды ровно 4 байта,

только команды загрузки-сохранения имеют доступ к оперативной памяти,

поддерживается три режима адресации (long, short и indexed),

декларируется 32 32-бит целочисленных, 32 64-бит вещественных и семь вспомогательных «теневых» (для обработки прерываний) регистров.

Вещественные регистры также могли быть адресованы как 64 32-бит или 16 128-бит. Использовалась 48-бит схема виртуальной адресации, расширенная с выходом спецификации PA-RISC 2.0 до 64-бит. Первые серийно выпускаемые ЦП уже соответствовали спецификации PA-RISC 1.1, наиболее существенным отличием от первой версии было использование раздельных шин и кэшей для команд и данных.

Первым ЦП архитектуры PA-RISC стал РА-7000, анонсированный в 1989 г. Его тактовая частота достигала 66 МГц при двухслойном 1,0-мкм технологическом процессе. Он располагал только одним пятистадийным целочисленным конвейером в составе Е-box, на который были также возложены функции А-box. Был предусмотрен (факультативно) и внешний математический сопроцессор. Подключались внешние асинхронные I-cache и D-cache размером до 256 Кбайт каждый, причем ширина шины к I-cache составляла 32 бит, а к D-cache — 64 бит. Каждый кэш был оснащен TLB, емкость которого равнялась 96 записям. Работа в многопроцессорном режиме предусмотрена не была.

Выпущенный в 1992 г. РА-7100 был уже суперскалярным RISC-процессором, располагающим, помимо одного целочисленного конвейера, одним вещественным. Соответственно отпала необходимость во внешнем математическом сопроцессоре. Благодаря переходу на трехслойный 800-нм процесс удалось поднять тактовые частоты до 100 МГц. Был увеличен и размер кэшей: I-cache до 1 Мбайт, D-cache до 2 Мбайт (оба на микросхемах асинхронной SRAM), а шина к I-cache была расширена до 64 бит, предусмотрен единый TLB на 120 записей и двухпроцессорный режим. Вскоре была выпущена незначительная модификация этого ЦП, РА-7150 (добавлена совместимость с cache bypass, а также внесены некоторые изменения в функциональной части ядра), тактовая частота которой была увеличена до 125 МГц.

В 1994 г. появился PA-7100LC (Low Cost, низкая цена), предназначавшийся для массовых систем. Это был однокристальный процессор со встроенным контроллером оперативной памяти FPM или EDO DRAM, со 128-бит шиной (с 16-бит каналом ЕСС). I-cache и D-cache были сведены в один унифицированный кэш объемом до 2 Мбайт, кроме того, был предусмотрен встроенный в ядро 1-Кбайт кэш команд. В структуру Е-box был добавлен второй конвейер. Количество записей в TLB сократилось до 64. Многопроцессорный режим предусмотрен не был. Для связи с периферией использовалась 32-бит шина GSC (General System Connect), причем данные и адреса передавались по одному и тому же каналу (мультиплексировались). Процессор изготавливался по трехслойному 75-нм технологическому процессу и был рассчитан на тактовые частоты до 100 МГц. Еще одним важным новшеством стало появление дополнительного набора команд МАХ-1, позволяющего проводить векторные расчеты, моделируя два 16-бит кванта данных в одном целочисленном регистре, что заметно ускоряло обработку однородных потоковых данных, например MPEG-кадров. Таким образом, Hewlett-Packard оказалась первой компанией, которая ввела поддержку SIMD-команд в архитектуру ЦП общего назначения.

На смену РА-7100 в 1995 г. пришел РА-7200. Он обладал всеми функциональными достоинствами РА-7100LC и был оснащен новой высокоскоростной системной шиной Runway, обеспечивающей скорость обмена данными до 960 Мбайт/с при частоте 120 МГц. Внешние I-cache и D-cache остались без изменений, но в ядро ЦП был встроен вспомогательный унифицированный 2-Кбайт кэш. ЦП мог быть использован в многопроцессорных системах — до четырех ЦП, подключенных к общей шине Runway. Добавилась поддержка порядка следования байт little-endian. Процессор изготавливался по 550-нм проектным нормам, тактовые частоты удалось довести до 140 МГц.

Появившийся в середине 1996 г. PA-7300LC оказался последним 32-бит ЦП этой архитектуры. Данный процессор представлял собой доработанный вариант PA-7100LC, в который благодаря четырехслойному 500-нм технологическому процессу удалось встроить 64-Кбайт I-cache и D-cache с двухканальной ассоциативностью. Был предусмотрен и внешний унифицированный B-cache с прямым отображением и сквозной записью, его объем мог достигать 8 Мбайт. Встроенный контроллер оперативной памяти был сохранен. Использовалась более скоростная модифицикация системной шины, GSC+. Как и у PA-7100LC, работа в многопроцессорном режиме не была предусмотрена. Хотя тактовые частоты возросли до 180 МГц, а показатели производительности этого ЦП были довольно хороши, лидером он не стал. Например, по результатам SPECint95 рабочая станция HP Visualize C160L, оборудованная 160-МГц PA-7300LC с 1-Мбайт B-cache, показывала производительность, близкую DEC AlphaStation 500/266 (266-МГц 21164 с 2-Мбайт B-cache), Sun Ultra 2 1200 (200-МГц UltraSPARC с 1-Мбайт B-cache) или Dell Dimension XPS Pro (200-МГц Pentium Pro с 256-Кбайт S-cache). По результатам SPECfp95 DEC AlphaStation 500/266 оказалась быстрее HP Visualize C160L на 47%, Sun Ultra 2 1200 — на 51%, a Dell Dimension

XPS Pro — медленнее на 16%. Таким образом, HP Visualize C160L вряд ли могла составить достойную конкуренцию 64-бит рабочим станциям на 21164 и UltraSPARC, а 32-бит рабочие станции на Pentium Pro находились в более низкой ценовой категории.

Необходимо отметить, что существовали и другие 32-бит ЦП архитектуры РА-RISC компании Hitachi. В 1993 г. появился РА/50 (изготавливаемый по трехслойному 600-нм техпроцессу) в двух вариантах:

PA/50L с тактовой частотой до 33 МГц и напряжением питания ядра 3,3 В предназначался для систем с низким энергопотреблением,

РА/50М с тактовой частотой до 66 МГц и напряжением питания ядра Б В — для стандартных конфигураций.

По архитектуре ядра обе модификации идентичны: только два конвейера (один целочисленный и один вещественный) и небольшие встроенные 8-Кбайт I-cache и 4-Кбайт D-cache — оба с двухканальной ассоциативностью. Традиционный для РА-7х00 большой скоростной внешний кэш отсутствовал. Имелся TLB I-cache на 32 записи и TLB D-cache на 64 записи. Был предусмотрен встроенный контроллер оперативной памяти с поддержкой SDRAM и аппаратной предвыборки данных (data prefetch).

Представленный в августе того же года на конференции Hot Chips процессор HARP-1 (Hitachi Advanced RISC Processor) обладал всеми функциональными достоинствами РА/50. Размер D-cache был увеличен до 16 Кбайт при сохранении двухканальной ассоциативности. Емкости обоих TLB были расширены до 128 записей каждый. Были добавлены внешние кэши команд и данных по 512 Кбайт каждый. Благодаря более совершенному четырехслойному 500-нм технологическому процессу, а также более высокому порогу энергопотребления, тактовые частоты ядра достигли 150 МГц.

Возвращаясь к продукции Hewlett-Packard, хотелось бы заметить, что к 1996 г. спрос на 32-бит RISC-серверы и рабочие станции упал, а конкуренты уже предлагали 64-бит решения. Поэтому не было ничего удивительного в том, что в январе вышла в свет уже 64-бит архитектура PA-RISC 2.0 и был представлен первый ЦП нового поколения — РА-8000.

Ядро ЦП было полностью перепроектировано, было реализовано внеочередное выполнение команд (предусмотрен буфер переупорядочения емкостью в 56 команд). Количество конвейеров ФУ увеличилось в несколько раз: четыре на Е-box (два основных и два вспомогательных), четыре на F-box (два для обработки операций умножения и сложения, два для расчета деления и вычисления квадратного корня), два на А-box (универсальные, т. е. каждый в состоянии рассчитывать виртуальные адреса и выполнять операции загрузки-сохранения). Учитывая ограничения пятислойного 500-нм техпроцесса, разработчики РА-8000 пошли на кардинальный шаг, вовсе отказавшись от интеграции кэш-памяти в ядро ЦП. А так как конвейеры ФУ остались довольно короткими (хотя мы и не располагаем информацией о точном количестве стадий каждого конвейера), то это фактически исключало возможность работы ЦП на высоких тактовых частотах. Это позволило использовать внешние I-cache и D-cache (оба двухпортовые и с прямым отображением, изготовленные из микросхем статической памяти), работающие на частоте, близкой к частоте ядра ЦП и подключенные к нему независимыми шинами. На практике использовались 6,7-нс синхронные 1-Мбит микросхемы SRAM с отложенной записью, формирующие кэши размером по 1 Мбайт. Был предусмотрен унифицированный полноассоциативный двухпортовый TLB на 96 записей, а также таблица истории условных переходов на 256 записей и кэш адресов переходов на 32 записи. Поддерживалось как статическое, так и динамическое предсказание условных переходов. В качестве системного интерфейса была применена уже опробованная с РА-7200 шина Runway. Также сохранилась работа в четырехканальном многопроцессорном режиме.

РА-8000 оказался вторым после POWER2 по количеству контактов на корпусе — 1085. Учитывая 64-бит целочисленную реализацию, был обновлен вспомогательный набор команд, который получил название МАХ-2 и предназначался для одновременной обработки уже четырех 16-бит квантов данных, находящихся в любом целочисленном регистре. Сохранилась впервые примененная в РА-7200 поддержка порядка следования байт little-endian, в дополнение к изначально заложенной в архитектуру big-endian. Важно, что при столь серьезных изменениях гарантировалась полная обратная совместимость с существующей 32-бит программной базой, написанной для предыдущих ЦП.

В мае 1997 г. была выпущена незначительно доработанная версия РА-8200, работающая на более высоких тактовых частотах, насколько это было возможно в рамках того же техпроцесса. Модификации подверглась логика предсказания условных переходов, отныне оперирующая таблицей истории переходов на 1024 записи. Емкость TLB была увеличена до 120 записей. Благодаря появлению на рынке 5-нс синхронных 4-Мбит микросхем SRAM с отложенной записью, объемы I-cache и D-cache увеличились вдвое. Заметим, что на момент выхода РА-8000 этот процессор обошел по производительности лидировавшего до этого 21164: при сравнении HP Visualize C180-XP (180-МГц РА-8000 с 1-Мбайт I-cache и 1-Мбайт D-cache) с DEC AlphaStation 5/333 (333-МГц 21164 с 2-Мбайт B-cache) первая оказалась производительнее в SPECint95 на 20%, а в SPECfp95 — на 50%. Лидерство РА-8000 продержалось всего несколько месяцев, пока стремительно набирающий тактовые частоты 21164А не вернул все на круги своя. 240-МГц РА-8200 смог вырваться вперед в SPECfp95, но уже не в SPECint95.

В сентябре 1998 г., благодаря переходу на пятислойный 250-нм техпроцесс, появился процессор РА-8500. Новый технологический процесс позволил достичь более высоких тактовых частот ядра ЦП, а также безболезненно интегрировать I-cache и D-cache, избавившись от сравнительно дорогих внешних микросхем SRAM. Размер I-cache составил 0,5 Мбайт с четырехканальной ассоциативностью, a D-cache — 1 Мбайт с четырехканальной ассоциативностью и обратной записью. В то же время задержки чтения из кэшей были небольшими, учитывая размеры кэшей: три такта. Размер таблицы истории переходов был увеличен в два раза (до 2048 записей), увеличилось и количество записей в TLB (со 120 до 160). Значительным изменениям подверглась системная шина Runway: она стала асинхронной (независимой от частоты ядра ЦП), а ее частота была зафиксирована на уровне 120 МГц при неизменной ширине канала данных в 64 бит (не считая битов ЕСС), эффективная частота удваивалась благодаря применению технологии DDR. На электрическом уровне шина также подверглась значительным изменениям: на смену 3,3-В терминируемой схеме push/pull пришла 1,5-В параллельно-терминируемая схема open drain. В итоге это позволило достичь эффективной скорости обмена данными по шине 1,92 Гбайт/с. Благодаря интеграции I-cache и D-cache количество контактов в разъеме ЦП сократилось почти вдвое, до 544.

В январе 2000 г. появилась модификация РА-8500 в рамках того же техпроцесса, работающая на более высоких тактовых частотах: РА-8600. В августе 2001 г. был выпущен РА-8700, уже по семислойному 180-нм техпроцессу с применением SOI — у него были увеличены до 0,75 Мбайт I-cache и до 1,5 Мбайт D-cache. Вдвое была увеличена емкость TLB (до 240 записей).

В 2003 г. после перехода на 130-нм технологию SOI был выпущен двухъядерный РА-88ОО с 0,75-Мбайт I-cache и 0,75-Мбайт D-cache в расчете на каждое ядро. Кроме того, был предусмотрен интерфейс для В-cache размером до 32 Мбайт. Системная шина была заменена на используемую ЦП семейства Itanium 2, что позволило максимально сблизить обе аппаратные платформы. Выпущенный в 2005 г. РА-8900 отличается лишь возросшими тактовыми частотами ядра, так как выпускается по тому же технологическому процессу, что и РА-8800. РА-8900 — это последний ЦП архитектуры РА-RISC, которая вскоре будет полностью замещена архитектурой Itanium. UNIX-подобная ОС для архитектуры РА-RISC несколько лет назад была перенесена на архитектуру Itanium со всей сопутствующей программной базой.

Необходимо отметить, что в отличие от активно развивавшегося поколения РА-7х00 в архитектуру ядра процессоров РА-8х00 не было внесено никаких серьезных изменений с момента их появления. Очевидно, что уже в то время руководство компании рассматривало вариант прекращения разработки этой архитектуры, и только затянувшийся запуск архитектуры Itanium позволил ей дожить до наших дней.

Все ЦП архитектуры РА-RISC, использующие системную шину Runway, позволяют создавать четырехканальные многопроцессорные системы. ЦП архитектуры РА-RISC использовались исключительно в серверах и рабочих станциях Hewlett-Packard и ее ближайших партнеров, таких, как Hitachi, Convex и Stratus. В свободную продажу они никогда не поступали.

К сожалению, компания Hewlett-Packard не пополняет базу данных SPEC2000 результатами систем с РА-8800 и РА-8900. Тем не менее, основываясь на показателях РА-8700 и приняв во внимание, что вдвое меньший размер D-cache, подключенного к каждому ядру, вполне компенсируется более высокой пропускной способностью системной шины, можно предположить, что производительность этих ЦП должна оцениваться 800—900 баллами как для SPECint2000, так и для SPECfp2000. Этого уже недостаточно для современных серверных ЦП. Добавление больших объемов кэш-памяти не приведет к значительному росту производительности, зато увеличит стоимость ЦП при производстве, а полная перепроектировка ядра требует значительных финансовых и временных затрат, поэтому решение Hewlett-Packard окончательно перейти на архитектуру Itanium выглядит вполне обоснованным.

**Sun SPARC**

Основы архитектуры SPARC (Scalable Processor ARChi-tecture, масштабируемая процессорная архитектура) были заложены в ходе работы над проектами RISC I и RISC II, проводимых в 1980-1982 гг. в университете Беркли (Калифорния, США) Дэвидом Паттерсоном (Dave Patterson), Манолисом Катевенисом (Manolis Katevenis) и Робертом Шерберном (Robert Sherburne). Компания Sun Microsystems, которая добилась того, что архитектура стала коммерчески успешной, была основана в феврале 1982 г. группой студентов, закончивших Стэнфордский университет: Скоттом МакНили (Scott McNealy), Вильямом Джоем (William Joy), Андреасом Бехтольсхаймом (Andreas Bechtolsheim) и Ви-нодом Хослой (Vinod Khosla).

Официальным началом разработки этой архитектуры принято считать момент, когда в 1984 г. Джой, вице-президент Sun по НИОКР, пригласил Паттерсона консультантом в компанию. Год спустя архитектура SPARC была представлена официально. Первая аппаратная реализация появилась в июле 1987 г. в виде рабочих станций Sun-4/260 и Sun-4/280, а вскоре Sun-4/ПО и Sun-4/150. В «железе» была реализована седьмая версия — SPARC V7. В то время архитектура SPARC была 32-бит, так как явной необходимости в 64-бит вычислениях и адресном пространстве тогда еще не наблюдалось (персональные компьютеры того времени еще только переходили на 32-бит вычисления).

Архитектура SPARC V7 обладала следующими особенностями:

длина любой команды составляет ровно 4 байта,

только команды загрузки-сохранения имеют доступ к оперативной памяти,

предусмотрен только один режим адресации (регистр+смещение на величину другого регистра или 13 бит

стековая регистровая модель (для любой программы доступно восемь основных 32-бит регистров и часть стека, состоящая из 24 вспомогательных 32-бит регистров), а также вещественные регистры.

Суть стековой регистровой модели архитектуры SPARC состоит в том, что для каждой запускаемой функции доступно 32 целочисленных регистра:

восемь глобальных регистров,

восемь входящих регистров (в которых находятся параметры, с которыми ьыла вызвана данная функция),

восемь регистров для локальных данных,

восемь исходящих регистров (в которых находятся

параметры, с которыми будет вызвана некоторая подфункция данной функции),

Очевидно, что те регистры, которые являются исходящими для данной функции, в то же время и входящие для ее некоторой подфункции. Другими словами, для каждой функции непосредственно выделяется только 16 вспомогательных регистров, а также обеспечивается доступ к восьми вспомогательным регистрам вышестоящей функции и восьми глобальным регистрам. Необходимые вспомогательные регистры выделяются функциями ОС из числа свободных. Преимущество такой регистровой модели — очень быстрое переключение задач (context switch), так как нет необходимости каждый раз сбрасывать данные из регистров в кэш или оперативную память, а потом считывать обратно в регистры, достаточно просто переместить регистровое «окно» в нужное место стека регистров.

В спецификации SPARC V8 документировано 32 32-бит вещественных регистра, которые могут быть при необходимости адресованы как 16 64-бит или 8 128-бит, предусмотрена возможность одновременного использования различной адресации. В SPARC V9 были регламентированы 16 дополнительных 64-бит регистров.

В 1991 г. вышла восьмая версия архитектуры SPARC (SPARC V8) — конструктивная доработка предыдущей, но все-таки еще 32-бит. Основными отличиями от 7-й версии было обязательное разделение кэша 1-го уровня на I-cache и D-cache, а также использование ранее отсутствовавших команд целочисленного умножения и деления. В мае 1992 г. Sun выпустила семейство рабочих станций SPARCstation 10, базирующихся на ЦП 8-й версии архитектуры, SuperSPARC (TMS390), совместно разработанном Sun и Texas Instruments. ЦП располагал I-cache размером в 20 Кбайт с пятиканальной ассоциативностью и D-cache размером 16 Кбайт с четырехканальной ассоциативностью, встроенным математическим сопроцессором и работал на тактовых частотах до 50 МГц. Кроме того, готовые системы обычно комплектовались B-cache. Сначала семейство SPARCstation 10 насчитывало четыре модели:

модель 30 с одним 36-МГц ЦП, без B-cache, с 32- Мбайт RAM и 420-Мбайт HDD за 18,5 тыс. долл. США;

модель 41 с одним 40-МГц ЦП, 1-Мбайт B-cache, 32-Мбайт RAM и 420-Мбайт HDD за 25 тыс. долл. США;

модель 52 с двумя 45-МГц ЦП, 1-Мбайт B-cache на каждый, 64-Мбайт RAM и 1-Гбайт HDD за 40 тыс.

модель 54 с четырьмя 45-МГц ЦП, 1-Мбай B-cache на каждый, 64-Мбайт RAM и 1-Гбайт HDD за 58 тыс. долл. США.

Вскоре было выпущено семейство рабочих станций SPARCstation 20, а также модернизирован появившийся в предыдущем году первый многопроцессорный сервер Sun — SPARCserver 600MP.

В 1992 г. вышла спецификация на версию 9 архитектуры (SPARC V9), которая должна была стать полностью 64-бит, но первых ЦП, удовлетворяющих ее требованиям, пришлось ждать еще несколько лет.

Первый ЦП с архитектурой SPARC V9 состоял из четырехстадийного одноконвейерного Е-box с одноконвейерным A-box?, a также из C-box с интерфейсом для 64-Кбайт внешнего унифицированного кэша с прямым отображением. Поскольку в ЦП не был предусмотрен F-box, использовался математический сопроцессор. Тактовая частота ЦП составляла 16 МГц (Sun4-260, Sun-4/280) или 14 МГц (Sun-4/ПО, Sun-4/150). Микросхема была создана совместно компаниями Sun и Fujitsu, изготавливалась на мощностях последней и маркировалась как МВ86900. Ими же была спроектирована микросхема контроллера математического сопроцессора (МВ86910), а сам сопроцессор состоял из двух микросхем — Weitek 1164 и 1165. Другими словами, первый SPARC состоял из четырех отдельных микросхем. Через некоторое время появился новый контроллер и новый сопроцессор к нему — Texas Instruments 8847. Они были применены в Sun-4/260 и Sun-4/280.

В октябре 1992 г. на Microprocessor Forum был официально представлен ЦП microSPARC — удешевленная версия SuperSPARC, также совместной разработки Sun и Texas Instruments. ЦП работал на тактовой частоте 50 МГц, располагал 4-Кбайт I-cache и 2-Кбайт D-cache, а самое главное — встроенным математическим сопроцессором. Этот ЦП предназначался для недорогих рабочих станций, которыми стали сразу же выпущенные SPARCclassic и SPARCstation LX по 4500 и 8000 тыс. долл. США соответственно. В 1993 г. появилась модификация этого ЦП, microSPARC II, с тактовой частотой от 60 МГц, 16-Кбайт I-cache и 8-Кб D-cache без B-cache. На его основе были выпущены популярные рабочие станции SPARCstation Voyager и SPARCstation 5.

В 1989 г. была представлена первая независимая разработка компании Cypress — ЦП 601. Впоследствии подразделение по созданию ЦП SPARC было выделено в дочернюю компанию Ross Technology, которая в мае 1992 г. анонсировала ЦП hyperSPARC. В мае 1993 г. Fujitsu приобрела контрольный пакет компании Cypress, а уже в августе того же года в продажу поступили 55- и 66-МГц версии hyperSPARC. В дальнейшем Ross предлагала разные версии hyperSPARC с тактовой частотой до 200 МГц и выше для модернизации рабочих станций и серверов компании Sun, в частности SPARCstation 10 и 20, а также SPARCserver 600МР. Еще одной оригинальной разработкой компании был SPARCplug — компактный модуль расширения, содержащий до двух ЦП hyperSPARC. ЦП Ross также применялись в некоторых моделях рабочих станций и серверов Sun. Однако к 1998 г. компания столкнулась с серьезными финансовыми проблемами, а после отказа Fujitsu от дальнейших капиталовложений руководству Ross пришлось распродавать активы и закрывать компанию.

Другой фирмой, работавшей над архитектурой SPARC, была Weitek. Кроме выпуска математических сопроцессоров для первых рабочих станций и серверов, в 1993 г. она представила SPARC PowerUP ЦП с тактовой частотой 80 МГц, предназначенный для модернизации Sun SPARCstation 2.

Еще одним важным участником рынка SPARC была HAL Computer Systems, основанная в 1990 г. бывшим инженером IBM Эндрю Хеллером (Andrew Heller), главным разработчиком архитектуры POWER. После согласия Fujitsu финансировать компанию была развернута разработка 64-бит SPARC64. Когда Хеллер ушел из компании (в конце 1993 г.), она стала структурой Fujitsu. В июне 1994 г. был получен первый работоспособный прототип SPARC64, а в феврале 1995 г. на ISSCC он был официально представлен, на три месяца опередив разработку самой Sun. Этот ЦП содержал беспрецедентно большие встроенные I-cache и D-cache — по 128 Кбайт каждый, B-cache отсутствовал. Следующий ЦП, SPARC64 GP, был выпущен в 1998 г. В нем был предусмотрен B-cache (от 2 до 8 Мбайт), процессор работал на более высоких тактовых частотах (до 810 МГц). В октябре 1999 г. на Microprocessor Forum главный разработчик SPARC64 V от HAL Computer Systems Майкл Шебанов (Michael Shebanow) открыл ключевые моменты архитектуры, несколько напоминавшей NetBurst: очень высокие тактовые частоты, кэш команд типа trace cache, очень быстрый 8-Кбайт D-cache. S-cache размером 512 Кбайт предполагалось разделить на сегменты для данных и команд, что было уникальным решением. Ядро ЦП получилось сравнительно большим — 65 млн. транзисторов. При использовании шестислойного 170-нм технологического процесса площадь ядра составила бы 380 мм2. Скорее всего, руководство Fujitsu сочло разработку экономически нецелесообразной, а потому в середине 2001 г. закрыло компанию HAL.

Sun UltraSPARC, 64-бит представитель девятой версии архитектуры, появился в мае 1995 г. Первые ЦП работали на тактовой частоте 143 и 167 МГц. Во втором полугодии к ним добавились версии с частотой 182 и 200 МГц. Е-box и F-box содержали по два девятистадийных конвейера, А-box — только один. Присутствовало 136 физических целочисленных регистров, 16-Кбайт I-cache с двухканальной ассоциативностью и 16-Кбайт D-cache с прямым отображением и сквозной записью. Были предусмотрены 4-Мбайт B-cache, работающий на частоте ядра ЦП, и унифицированный полноассоциативный TLB на 64 записи. Также ЦП располагал блоком предсказания переходов, управляющей таблицей истории переходов на 2048 записей. Процессор работал с 44-бит виртуальным и 41-бит физическим адресным пространством. Был внедрен набор вспомогательных векторных инструкций VIS (Visual Instruction Set), обрабатывавшихся на обоих вещественных конвейерах. Добавилась также поддержка порядка следования байтов little-endian. ЦП изготавливался по четырехслойному 500-нм техпроцессу на мощностях Texas Instruments. Примечательно, что этот ЦП не имел средств внеочередного выполнения команд (забегая вперед, отметим, что это относится ко всем ЦП семейства UltraSPARC и по сей день).

В сентябре 1996 г. Fujitsu объявила о выпуске 32-бит TurboSPARC (MB86907) с тактовой частотой 170 МГц, предназначенного как для модернизации существующих рабочих станций (например, Sun SPARCstation 5), так и для выпуска новых. ЦП содержал I-cache и D-cache размером по 16 Кбайт, поддерживал до 1 Мбайт B-cache, а также отличался низким энергопотреблением (около 7 Вт). Этого удалось добиться благодаря использованию передового 350-пм технологического процесса компании Fujitsu. На момент начала продаж цена составляла 500 долл. США.

Еще в октябре 1995 г. на Microprocessor Forum был анонсирован UltraSPARC II. Тем не менее выпуск этого ЦП затянулся до января 1997 г. Не претерпев почти никаких изменений в функциональной части, этот ЦП, по сути, был модификацией UltraSPARC, изготовленной по новым 350-нм проектным нормам компанией Texas Instruments, что позволило довести тактовые частоты до 300 МГц, а к маю 1998 г. до 360 МГц. Объем B-cache был увеличен до 16 Мбайт. После перехода Texas Instruments на 250-нм техпроцесс и перепроектировки ядра ЦП под его требования тактовые частоты были повышены до 450 МГц в ноябре того же года.

В январе 1998 г. была выпущена модифицированная версия UltraSPARC II — UltraSPARC Hi. В ядро ЦП встроили контроллер оперативной памяти (50-нс EDO DRAM с 64-бит шиной данных и 8-бит каналом ЕСС) и шины PCI 2.1. Процессор выпускался по 350-нм проектным нормам (пять слоев), что позволило достичь тактовой частоты 360 МГц, а с переходом на 250-нм поднять до 480 МГц. В 2000 г. с переходом на 180-нм технологию с алюминиевыми проводниками в ядро ЦП был впервые встроен S-cache (256 Кбайт), B-cache был упразднен, а контроллер оперативной памяти модифицирован для работы с 100-МГц SDRAM. Тактовая частота ЦП достигла 500 МГц. После перехода в 2002 г. на техпроцесс с тем же уровнем детализации и применением медных проводников объем S-cache был увеличен до 512 Кбайт, а тактовые частоты ядра ЦП до 650 МГц. Выпущенный в сентябре 2000 г., UltraSPARC III имел переработанное ядро UltraSPARC. Количество стадий целочисленных конвейеров возросло до 14, вещественных — до 16, число целочисленных конвейеров — до трех (один из конвейеров мог обрабатывать и команды загрузки). Серьезным изменениям подверглась подсистема кэш-памяти: I-cache был увеличен до 32 Кбайт (четырехканальная ассоциативность), а D-cache — до 64 Кбайт (четырехканальная ассоциативность, обратная запись). B-cache объемом до 8 Мбайт состоял из 5-нс микросхем синхронной SRAM с 256-бит каналом данных, причем теги B-cache хранились в ядре ЦП. Для оптимизации работы с B-cache были предусмотрены встроенный 2-Кбайт кэш предварительной выборки и 2-Кбайт кэш записи с четырехканальной ассоциативностью. Встроенный контроллер оперативной памяти SDRAM имел 150-МГц 128-бит шину данных. Ширина системной шины также была 128 бит. ЦП обеспечивал 64-бит виртуальную адресацию и 43-бит физическую. В многопроцессорных конфигурациях применялась топология общей шины (Sun Fireplane, 150 МГц), к которой подключалось до четырех ЦП. Процессор изготавливался по 180-нм проектным нормам, что позволило достичь тактовой частоты 900 МГц, а после перехода на семислойный 130-нм процесс — 1200 МГц. Благодаря широким внешним интерфейсам этот ЦП занял второе (после POWER2) место по числу выходных контактов — 1368.

В октябре 2002 г. на Microprocessor Forum был представлен Fujitsu SPARC64 V. У него совершенно новый дизайн, отличный от одноименного детища Fujitsu HAL Computer Systems. Дело в том, что после поглощения Amdahl Computer в апреле 2002 г. Fujitsu начала разработку собственного SPARC64 V на базе HAL SPARC64 GP.

Fujitsu SPARC64 V довольно сильно отличался от семейства ЦП UltraSPARC, он поддерживал внеочередное выполнение команд, что предполагало совершенно иную функциональную организацию ядра. ЦП имел 11-стадийный двухконвейерный Е-box, 12-стадийный двухконвейерный F-box и 15-стадийный двухконвейерный А-box. Для обеспечения эффективного внеочередного выполнения каждый из конвейеров имел собственный буфер команд (reservation station). Был предусмотрен вспомогательный набор команд VIS. Процессор имел два файла целочисленных регистров: большой архитектурный (160 регистров с одним портом чтения и двумя портами записи) и стандартный основной (32 регистра с восемью портами чтения и двумя портами записи). При переключении задач или вызове подфункции содержание 16 регистров из основного файла копировалось в соответствующие регистры архитектурного файла, а оттуда считывались нужные регистры для новой задачи. Объемы I-cache и D-cache остались теми же, что и у самого первого SPARC64 (по 128 Кбайт), оба с двухканальной ассоциативностью и 64-бит размером строки, a D-cache к тому же был двухпортовым и с возможностью обратной записи. Был добавлен 2-Мбайт S-cache с четырехканальной ассоциативностью и обратной записью, с 64-бит строкой, причем ширина канала данных к нему составляла 256 бит. Для работы с S-cache было предусмотрено два промежуточных буфера: входящий на 16 строк и исходящий на 8. Примечательно, что как S-cache, так и D-cache обладали логикой ЕСС (обычно D-cachc защищается только контролем четности), это отличие было унаследовано от SPARC64 GP. Необычной была и конструкция TLB: для S-cache (страницы данных) было предусмотрено по 1024 записи с двухканальной ассоциативностью для 8- и 4-Кбайт страниц, а для D-cache — 32 полноассоциативных записи для страниц любых размеров. Точно такая же организация была и у TLB команд.

Логика предсказания условных переходов оперировала таблицей истории переходов с четырехканальной го контроллера памяти не было, системная шина (со звездообразной топологией) со 128-бит каналом данных и применением технологии DDR обеспечивала пропускную способность 3,52 Гбайт/с. Процессор работал с 64-бит виртуальным адресным пространством и 43-бит реальным. ЦП изготавливался по 130-нм технологии SOI (восемь слоев) с медными проводниками, тактовая частота составила 1350 МГц, а после перехода на десятислойную 90-нм SOI-технологию увеличилась до 2 ГГц (одновременно объем S-cache был увеличен до 4 Мбайт).

В 2003 г. была выпущена модификация UltraSPARC III, оборудованная 1-Мбайт S-cache с четырехканальной ассоциативностью и обратной записью — С IIIi. Встроенный контроллер оперативной памяти был также модернизирован и позволил использовать DDR SDRAM. Был реализован новый тип системной шины (JBus, от 150 до 200 МГц) и внедрен контроллер шины PCI. Процессор изготавливался по 130-пм технологии (семь слоев с медными проводниками), а впоследствии и по 90-нм. Тактовые частоты ядра ЦП — до 1600 МГц.

В феврале 2004 г. появился двухъядерный вариант UltraSPARC III — UltraSPARC IV. Ядра существенно не изменились, но контроллер оперативной памяти, системный интерфейс и контроллер B-cache были общими для двух ядер, что необходимо для сохранения совместимости с UltraSPARC III. Изначально изготавливаемый по 130-нм технологическому процессу ЦП работал на тактовой частоте до 1200 МГц. После перехода на 90-нм процесс тактовая частота возросла до 1350 МГц.

Будущее семейства UltraSPARC туманно — в апреле 2004 г. разработка UltraSPARC V была остановлена, а большая часть его разработчиков уволена. Не ясны и перспективы ожидаемого в конце этого года двухъядерного UltraSPARC IV+ (Panther).

Что касается архитектуры SPARC в целом, то ее будущее связано с двумя примечательными проектами в области аппаратных средств — процессорами нового поколения Niagara и Rock.

Процессор Niagara и системы на его базе должны быть представлены в конце 2005, либо в самом начале 2006 года. На практике он выглядит следующим образом — 90 нм технология производства, восемь полноценных самостоятельных ядер на кристалле, каждое ядро способно исполнять до четырех вычислительных потоков одновременно. Итого в одном кристалле пользователям доступно 32 «виртуальных» процессора, способных выполнять возложенные на них задачи не хуже реальных физических устройств. Высокая степень интеграции позволила разместить на самом кристалле контроллер памяти стандарта DDR2 и четыре контроллера Gigabit Ethernet. По предварительным данным первые образцы будут функционировать на частоте 1,2 ГГц при тепловыделении на уровне примерно 50 Вт. В дальнейшем предполагается значительно поднять тактового частоту процессора. Проект Niagara принципиально одно процессорное решение. В значительной степени это связано с желанием Sun отработать на этом этапе заложенные в основу Niagara принципы организации еще и межпроцессорных взаимодействий. В следующем поколении, известном под условным именем Niagara 2, предполагается реализовать необходимые для построения SMP-систем компоненты. Niagara в некотором роде специализированный процессор — он разрабатывался с учетом результатов исследований об особенностях работы многопоточных приложений и наилучшим образом подходит для построения серверных систем, ориентированных на предоставление сетевых сервисов. Стоит отметить, что по информации предоставленной Sun, процессор Niagara будет ориентирован на использование в массовых системах нижнего ценового уровня, цена которых будет сравнима или даже ниже аналогичных по функциональности систем, построенных на базе платформы х86.

Микропроцессор Rock, появление которого намечено на 2008 год, представляет собой дальнейшее развитие идей, реализованных в проекте Niagara, однако в этой реализации «центр тяжести» перенесен на построение серверных систем занятых интенсивной обработкой данных, для чего в архитектуре сделаны соответствующие изменения.

**MIPS**

MIPS — старейшая из существующих сегодня RISC-архитектур. Ее основы были заложены в 1981 г. во время работы над одноименным проектом в Стэнфордском университете, руководителем которого был Джон Хеннесси (John Hennessy). MIPS — это аббревиатура от Microprocessor without Interlocked Pipeline Stages, т. е. микропроцессор без блокированных стадий конвейеров. Суть архитектуры следует из названия — в ней реализована идея оптимизации работы конвейеров благодаря упрощению набора команд.

Современные ЦП выполняют команды самого разного назначения, причем многие из них просты и могут быть выполнены за один такт. Но некоторые команды, например, целочисленное умножение или деление, а также часть вещественных, например вычисление квадратного корня, требуют большого числа процессорных тактов. Когда такая команда выполняется, логика планировщика ЦП должна приостановить (блокировать) подачу на конвейер других команд. Разработчики архитектуры MIPS решили исключить из логики ЦП все команды, которые не могли быть выполнены за один такт.

Это, разумеется, не означает, что ЦП архитектуры MIPS не в состоянии выполнять сложные операции, просто расчеты будут выполняться при помощи серии более простых команд. Это также не означает, что такие ЦП окажутся малоэффективными и непроизводительными: например, архитектура Alpha тоже не предусматривала аппаратное выполнение целочисленного деления, но ЦП этой архитектуры долгое время были эталоном производительности. Пожертвовав быстродействием на некоторых операциях, разработчики получили простое и легко синхронизируемое ядро, предоставляющее возможность добиться высоких тактовых частот, увеличить количество конвейеров ФУ, реализовать более мощные блоки предсказания переходов и увеличить объемы интегрируемой кэш-памяти.

В 1984 г. Хеннесси, уверенный в коммерческом успехе своего мероприятия, уходит из Стенфорда и основывает MIPS Computer Systems. Первым MIPS-процессором стал выпущенный в начале 1986 г. 32-бит R2000 с тактовой частотой 8 МГц, состоящий из 110 тыс. транзисторов и изготавливаемый по 2-мкм технологическому процессу. Логика F-box отсутствовала (что неудивительно, учитывая количество транзисторов в ядре ЦП), к ЦП подключался математический сопроцессор R2010.

В 1988 г. вышел доработанный вариант процессора R3000, работающий в многопроцессорном режиме при почти том же количестве транзисторов. Он также дополнялся математическим сопроцессором R3010.

В феврале 1991 г. MIPS анонсировала свой ЦП следующего поколения — 64-бит R4000 со встроенным математическим сопроцессором. Его тактовая частота составила 100 МГц, первые партии ЦП были отгружены в марте 1992 г. Таким образом, именно MIPS стала первой 64-бит архитектурой, опередив DEC Alpha. Для обеспечения требуемой нагрузки на конвейеры ФУ при работе на столь высоких тактовых частотах в ядро были впервые встроены 8-Кбайт I-cache и D-cache, предусмотрен внешний B-cache объемом до 1 Мбайт. ЦП получился далеко не таким доступным по цене, как R2000 и R3000 — реализация 64-бит команд и встроенного кэша потребовала 1,35 млн. транзисторов, а площадь ядра составила 213 мм2.

Длина целочисленного конвейера была увеличена до восьми стадий, тогда это считалось очень длинным конвейером. ЦП изготавливался по весьма совершенному в то время 800-нм процессу и обладал максимальной рассеиваемой мощностью 15 Вт (для сравнения: у R3000 значение соответствующей характеристики 4 Вт). В мае 1993 г. компания объявила о выпуске R4400, доработанного варианта R4000. Он отличался вдвое увеличенными I-cache и D-cache, а также исправлением ошибок 64-бит режима. Количество транзисторов возросло до 2,3 млн.

К началу 1990-х гг. у компании начались финансовые проблемы, и R4000 мог оказаться последней разработкой компании. В марте 1992 г. Silicon Graphics — основной потребитель ЦП MIPS — объявила о своем намерении купить компанию. MIPS Computer Systems стала дочерней компанией Silicon Graphics и была переименована в MIPS Technologies. С этого момента компания занимается исключительно исследованиями, разработкой ЦП и лицензированием своих разработок.

В июне 1994 г. MIPS анонсирует свой первый ЦП с целочисленной суперскалярностью — R8000, который располагал двухконвейерными Е-box и А-box, 16-Кбайт I-cache с прямым отображением и двухпортовым 16-Кбайт D-cache с прямым отображением и сквозной записью. Также присутствовал кэш условных переходов на 1024 записи. Двухконвейерный F-box был вынесен во внешний сопроцессор (R8010). Ни I-cache, ни D-cache не поддерживали никаких алгоритмов определения и коррекции ошибок. Подсистема S-cache состояла из микросхемы контроллера, двух логических банков S-тегов и двух логических банков синхронного S-cache (в обоих случаях по одному на четные и нечетные адреса). В итоге объем S-cache с четырехканальной ассоциативностью и обратной записью мог достигать 16 Мбайт. Двухпортовый трехканальный унифицированный TLB позволял хранить 384 записи.

Оба целочисленных конвейера имели по одному вычислительному блоку для простых команд, а также блок смещений и блок умножения-деления, которые были общими для обоих конвейеров. Файл целочисленных регистров располагал девятью портами чтения и четырьмя портами записи. R8000, R8010 и контроллер S-cache использовали общую 80-бит шину TBus, работающую на частоте ЦП. Системная шина ЦП со 128-бит каналом данных и 40-бит каналом адресов также работала на частоте ЦП.

R8000 и R8010 состояли из 2,8 и 0,8 млн. транзисторов соответственно, изготавливались по 500-нм проектным нормам (три слоя), что позволило достичь тактовой частоты 75МГц, которая год спустя была повышена до 90 МГц. Несмотря на столь невысокие тактовые частоты, R8000 и R8010 демонстрировали сравнительно неплохую производительность. Например, рабочая станция Silicon Graphics Power Indigo2 (75-МГц R8000 и R8010 с 2-Мбайт B-cache) по результатам SPECfp92 не уступала DEC 3000 Model 900 (275-МГц 21064А с 2-Мбайт B-cache), а по SPECint92 — была на 47% медленнее.

В марте 1996 г. компания выпускает R10000, анонсированный еще в октябре 1994 г. Основываясь на R4000, но учитывая наработки R8000 и преимущества внедрения внеочередного выполнения команд с переименованием регистров, компания создала ЦП, послуживший основой для всех ее последующих разработок вплоть до наших дней. Все поступающие в ЦП команды (по четыре за такт) немедленно проходили процедуру предварительного декодирования и сортировки, в ходе которой к ним добавлялось по четыре дополнительных байта, и в таком виде они помещались в 32-Кбайт 1-cache с двухканальной ассоциативностью. Для предсказания переходов использовалась таблица истории, состоящая из 512 двухбитовых записей. I-box проводил выборку и декодирование команд из I-cache со скоростью четыре команды за такт, после чего они направлялись в очереди емкостью по 16 команд, по одной очереди на Е-box (пять стадий, два конвейера), F-box (семь стадий, два конвейера) и А-box (один шестистадийный конвейер).

Для успешной работы алгоритмов внеочередного выполнения и переименования регистров имелось два файла регистров по 64 записи в каждом. Файл целочисленных регистров располагал 12 портами чтения и четырьмя портами записи, а файл вещественных регистров — 16 и четыре соответственно. Оба целочисленных конвейера полнофункциональны, но только первый может выполнять команды смещения, а второй — умножения и деления. Первый вещественный конвейер рассчитан на простые арифметические операции, а второй — на команды умножения, деления и извлечения квадратного корня. Почти все целочисленные команды (кроме умножения и деления) выполнялись за один такт, а почти все вещественные (кроме умножения, деления и вычисления квадратного корня) — за три такта. Выборка данных производилась из 32-Кбайт D-cache с двухканальной ассоциативностью, предусмотрен также 16-Мбайт B-cache с двухканальной ассоциативностью и 128-бит каналом данных, работающего со скоростью от 1/3 до полной частоты ядра ЦП.

Емкость унифицированного TLB — 64 записи. Также была реализована новая 64-бит системная шина (Avalanche), работающая со скоростью от 1/4 до полной частоты ядра ЦП, с мультиплексированием каналов данных и адресов, выполняющая до восьми одновременных операций.

Был предусмотрен многопроцессорный режим — топология общей шины, до четырех ЦП. R10000 совместим с набором команд MIPS V, используемым для векторизации вещественных расчетов путем одновременной обработки двух 32-бит чисел одинарной точности в любом вещественном регистре, а также MDMX (MIPS Digital Media Extension, аналог Intel MMX) для векторизации целочисленных вычислений, с позиционированием новых регистров на существующие вещественные.

Процессор изготавливался по 350-нм технологии на мощностях NEC и Toshiba, содержал 0,8 млн. транзисторов и работал с тактовой частотой до 180 МГц.

В ноябре 1998 г. вышла модификация R12000, перепроектированная под 250-нм техпроцесс. Размер таблицы истории переходов был увеличен до 2048 записей, некоторые изменения были внесены также в логику I-box и С-box. Количество транзисторов возросло до 7,15 млн., тактовые частоты были подняты до 300 МГц и выше. В дальнейшем производство перешло на 180-нм техпроцессы.

В 2001 г. был выпущен процессор R14000 — практически R12000, перепроектированный под 130-нм проектные нормы, с увеличенной до 600 МГц тактовой частотой. Выпущенный в 2002 г. R16000 отличался лишь удвоенными размерами I-cache и D-cache, а также возросшей до 700 МГц тактовой частотой вследствие перехода на 110-нм проектные нормы.

В настоящее время MIPS Technologies — уже не дочерняя компания Silicon Graphics. После катастрофического падения объемов продаж рабочих станций и серверов на архитектуре MIPS в конце 1990-х гг. Silicon Graphics пришлось обратить внимание на более производительные и перспективные архитектуры. Silicon Graphics предлагает спектр серверных решений на основе архитектуры Itanium (модельный ряд Altix), а выпуск продукции архитектуры MIPS постепенно сворачивается. В 1998-2000 гг. MIPS Technologies получила свободу действий, а на сегодня ее источником дохода являются лицензионные отчисления от все еще выпускаемых ЦП архитектуры MIPS (преимущественно 32-бит встраиваемых), а также прибыль от контрактных проектировочных работ. Сегодня ЦП архитектуры MIPS можно встретить в самых неожиданных местах: от серверов и рабочих станций до сетевых маршрутизаторов (преимущественно компании Cisco Systems) и игровых приставок.

**IBM POWER**

История архитектуры POWER (Performance Optimisation With Enhanced RISC, оптимизация производительности с расширенным RISC) началась в 1970-х гг., когда Джон Кок (John Соске) из Исследовательского центра Томаса Дж. Ватсона возглавил работу над проектом по разработке самого первого RISC ЦП (хотя собственно термины RISC и CISC происходят из Университета Беркли). Изначально специалисты просто изучали работу мэйнфреймов IBM типа System/370 с целью их оптимизации. Но в ходе исследований появилась идея о создании ЦП с компактным ядром, способным работать на высоких тактовых частотах. Были начаты работы в этом направлении, и в 1977 г. появился IBM 801. Этот ЦП применялся во многих типах компьютерной техники IBM как универсальный вспомогательный контроллер, а после основательной доработки дал начало новой архитектуре.

В 1990 г. IBM начала поставки рабочих станций семейства RS/6000 (RISC System/6000), в основе которых 32-бит ЦП POWER1 с тактовой частотой от 20 до 30 МГц. Этот суперскалярный ЦП состоял из трех микросхем: процессора целочисленной логики с 8-Кбайт I-cache, математического сопроцессора и блока предсказания условных переходов (в сумме около 0,8 млн. транзисторов). Был предусмотрен внешний D-cache объемом 32 или 64 Кбайт. Набор команд этого ЦП содержал свыше 100 инструкций, что было довольно много для представителя подхода RISC, поэтому иногда архитектуру POWER считают чем-то средним между «настоящими» RISC и «классическими» CISC.

Этот ЦП дал начало семейству PowerPC, разрабатываемому альянсом IBM, Motorola и Apple Computer. ЦП этого семейства применялись во многих персональных компьютерах, рабочих станциях (Apple Power Macintosh или PowerMac) и даже в некоторых серверах. Далее речь пойдет о разработках IBM для ее собственных нужд, т. е. ЦП семейства POWER, хотя родственность архитектур POWER и PowerPC сохраняется до сих пор.

В 1993 г. был выпущен POWER2. Примечательно, что рабочие станции и серверы на его основе продавались вплоть до 1998 г. Это лишний раз свидетельствует о том, что инженеры IRM создали удачную и сбалансированную архитектуру. POWER2 содержал два целочисленных конвейера, два вещественных и два для обработки условных и безусловных переходов. Физически POWER2 состоял из восьми микросхем в одном общем корпусе):

ICU (Instruction Cache Unit) — I-cache, I-box и логика обработки переходов,

FXU (FiXed-Point Unit) — E-box,

FPU (Floating-Point Unit) — F-box,

четыре DCUs (Data Cache Units) — D-cache, а также контроллер линий данных шины памяти,

SCU (Storage Control Unit) — контроллер управляющих и адресных линий шины памяти, а также периферийной шины.

Связь между ICU, FXU и SCU обеспечивала скоростная внутрипроцессорная шина, ширина канала данных шины памяти составляла 256 бит (если точнее, то по 64 бита данных и 8 бит для ЕСС на каждый DCU). Контроллер оперативной памяти был интегрирован в ЦП. Ширина связующего канала между четырьмя DCU и I-box составляла 128 бит, Е-box — 64 бита, F-box — 256 бит. I-box, E-box и F-box содержали по два конвейера. Объем I-cache с двухканальной ассоциативностью составлял 32 Кбайт, а трехпортового D-cache с четырех-канальной ассоциативностью и обратной записью — 256 Кбайт (по 64 Кбайт в каждом DCU). D-cache позволял одновременно выполнять до двух операций со стороны ФУ (чтение или запись) и одну операцию со стороны оперативной памяти. DCU работали на утроенной частоте ЦП. ICU производил выборку восьми команд за такт из I-cache и размещал их в очереди FXU и FPU (емкостью по восемь команд каждая), кроме тех команд, которые обрабатывал сам. FXU содержал два синхронизирующихся файла целочисленных регистров, по одному на каждый конвейер, с четырьмя портами чтения и четырьмя портами записи у каждого. Файлы состояли из 32 регистров по 32 бит. Целочисленные конвейеры были неоднородны — только один из них мог выполнять команды умножения и деления, но оба выполняли функции адресных.

Блок TLB команд с двухканальной ассоциативностью содержал 128 записей, a TLB данных — 512 (ассоциативность также двухканальная). FPU поддерживал технику переименования регистров, а поэтому в его файле их было предусмотрено 54 (вместо 32, согласно спецификации на архитектуру) 64-бит регистра. Связь с периферийными устройствами обеспечивал блок SCU, для этого использовалась 10-МГц шина MCA (Micro-Channel Architecture). В итоге изготавливаемый по 720-нм проектным нормам ЦП состоял из 23,04 млн. транзисторов и имел суммарную площадь 1217 мм2. Первые модификации работали на тактовой частоте от 55 до 71 МГц. Количество контактов на корпусе ЦП побило все существующие рекорды — 3181.

В 1996 г. инженеры IBM модифицировали POWER2. Так как к тому времени был доступен пятислойный 290-им техпроцесс, гораздо более совершенный по сравнению с четырехслойным 450-нм, использовавшимся поначалу, то появилась возможность реализовать все блоки в одной микросхеме — это позволяло увеличить максимальную тактовую частоту. Этот ЦП получил наименование P2SC (POWER2 Single Chip). Из изменений в функциональной части можно отметить вдвое уменьшенные размеры D-cache и TLB данных (очевидно, вынужденный шаг, ведь возможности нового техпроцесса были отнюдь не безграничны). Так как ЦП был выпущен на рынок с тактовой частотой 135 МГц, то частоту шины памяти пришлось зафиксировать на 1/2 частоты ЦП. Количество транзисторов составило 15 млн., из них 9,3 млн. приходилось на I-cache и D-cache. Площадь ядра 335 мм2, т. е. значительно меньше, чем оригинального POWER2, а количество контактов сократилось до 1088. P2SC использовался при постройке суперкомпьютера IBM SP2, содержащего до 512 ЦП и до 1 Тбайт оперативной памяти.

Так как ядро POWER2 не смогло достичь тактовой частоты выше 160 МГц, то при разработке POWER3 за основу было взято ядро массового PowerPC 620. Существенных изменений в него не вносилось, так как производительность PowerPC 620 была достаточно высокой. В ядро добавили по одному вещественному и адресному конвейеру, увеличили количество регистров переименования и очередей команд. ЦП мог работать с обоими наборами команд — PowerPC и POWER, a целочисленная и адресная логика была расширена до 64-бит. Кроме того, как и PowerPC 620, этот ЦП допускал внеочередное выполнение команд.

ЦП располагал восемью конвейерами: тремя целочисленными (двумя простыми и одним комплексным), двумя вещественными, двумя адресными и одним для логики переходов. Было предусмотрено два синхронизируемых файла целочисленных регистров (по 48 записей), а также файл вещественных регистров на 56 записей. Объем I-cache со 128-канальной ассоциативностью составлял 32 Кбайт, a D-cache со 128-канальной ассоциативностью и обратной записью — 64 Кбайт. TLB команд и данных хранили по 128 записей, также присутствовала таблица истории переходов на 2048 записей и кэш адресов переходов на 256 записей. B-cache, работающий на тактовой частоте 200 МГц, подключался через выделенный 256-бит канал данных. Ширина системной шины составила 128 бит, а тактовая частота — 100 МГц. ЦП изготавливался по гибридному технологическому процессу с 250-нм транзисторами и 350-нм слоями металлизации и содержал 15 млн. транзисторов (площадь составляла 270 мм2). Первые ЦП работали на частоте 200 МГц, но с переходом на 180-нм техпроцесс SO1 с медными проводниками она была существенно повышена.

Первые новости о POWER4 разошлись по миру с Microprocessor Forum, который проходил в октябре 1999 г. Вслед за объявлением о, скором выходе Merced, сделанным Intel, компания IBM намекнула о явном нежелании переходить на архитектуру VLIW в обозримом будущем и анонсировала свой двухъядерный POWER4. По сути, этот ЦП представлял собой два модифицированных ядра POWER3, универсальный коммутатор, 1,44 Мбайт восьмиканального многопортового S-cache с обратной записью и вспомогательные устройства, собранные на одной кремниевой подложке. Объем I-cache с двухканальной ассоциативностью составил 64 Кбайт, а

D-cache с четырехканальной ассоциативностью и обратной записью — 32 Кбайт. D-cache был трехпортовым и позволял выполнить одновременно две операции чтения и одну — записи. Объем B-cache с восьмиканальной ассоциативностью и обратной записью мог достигать 32 Мбайт, а его теги хранились непосредственно на подложке ЦП (для гораздо более быстрого доступа). По четыре таких ЦП объединялись в один процессорный модуль. Оба ядра разделяли между собой целую серию шин:

две 128-бит однонаправленные к B-cache и далее к оперативной памяти (работавшие на l/3 тактовой частоты ЦП),

шесть 128-бит однонаправленных для связи с двумя соседними ЦП в пределах своего процессорного модуля (работавшие на 1/2 тактовой частоты ЦП),

две 64-бит однонаправленные для связи с ЦП в соседних модулях,

две 32-бит однонаправленные для связи с периферийными контроллерами.

В пределах модуля процессоры соединялись между собой по кольцевой топологии, причем размещались на пластине модуля под углом в 90° друг к другу.

Таким образом, в одном процессорном модуле содержалось четыре ЦП (восемь ядер), а также четыре микросхемы B-cache. Непосредственно могло быть связано четыре процессорных модуля, вместе составлявших основу системы с 16 двухъядерными ЦП. Выпущенный в 2001 г. ЦП изготавливался по 180-нм технологии SOI с медными проводниками, состоял из 174 млн. транзисторов и работал на тактовой частоте в 1 ГГц, а позже были представлены и более скоростные версии. Энергопотребление одного процессорного модуля превышало 1 кВт. Впоследствии был выпущен POWER4+ — модификация процессора, изготовленная по 130-нм проектным нормам SOI (восьмислойный), содержавшая уже 184 млн. транзисторов при площади ядра ЦП в 267 мм2 и работавшая на тактовой частоте до 1,9 ГГц.

В октябре 2003 г. на Microprocessor Forum был представлен POWER5, доработанный вариант POWER4. Наиболее важным отличием была внутриядерная многопоточность, которая моделировала два логических ядра из одного физического методами, несколько напоминающими Intel Hyper-Threading. Объем S-cache был увеличен до 1,92 Мбайт, а его ассоциативность — до десяти каналов. Был существенно модифицирован и внутри-процессорный коммутатор, благодаря чему все внутри-модульные шины стали работать на полной частоте ЦП, а междумодульные — на 1/2 частоты ЦП. Встроенный контроллер памяти стал совместим не только с DDR SDRAM, но и с DDR2 SDRAM.

POWER5 изготавливался по тому же техпроцессу, что и POWER4+, содержал 276 млн. транзисторов при общей площади ядра ЦП 389 мм2 и работал на тактовой частоте до 1,9 ГГц. Как и POWER4 с POWER4+ в свое время, POWER5 стал процессором с самым большим количеством контактов — 5370.

Подводя итоги, хотелось бы заметить, что у IBM есть действительно отличный серверный и суперкомпьютерный ЦП, пригодный для широкого круга задач. Сожаление может вызывать лишь нежелание компании IBM поставлять его независимым сборщикам и интеграторам, ограничившись продвижением массового PowerPC 970.

**Заключение**

В начале 1990-х гг. рынок оказался фактически поделен между компьютерной техникой на основе ЦП х86-архитектуры и производительными, но дорогими рабочими станциями на различных RISC-архитектурах. По результатам продаж за 1992-й г. рынок рабочих станций выглядел так:

РА-RISC — 31%;

SPARC — 25%;

MIPS — 20%;

POWER — 12%;

остальные — 12%.

С ростом популярности и производительности решений на основе архитектуры Intel x86 параллельно со снижением их цены сложилась ситуация, в которой доля рынка рабочих станций на основе «настоящих» RISC-архитектур сократилась до нерентабельного минимума, что привело к окончательному смещению приоритетов разработчиков этих архитектур в сторону серверов и суперкомпьютеров. Как видим, многие из этих RISC-архитектур уже остановились в своем развитии и их исчезновение — вопрос времени. Прекратилась разработка DEC Alpha и HP PA-RISC, будущее MIPS на рынке серверов и рабочих станций весьма сомнительно, IBM POWER (если не принимать во внимание PowerPC) позиционируется почти исключительно на рынок дорогих высокопроизводительных серверов, a Sun SPARC явно переживает не лучшие времена. Пожалуй, наилучшие перспективы у Itanium 2, разрабатываемого совместно Intel и HP.

Совсем недавно Sun приоткрыла завесу тайны над проектами процессоров нового поколения с кодовыми наименованиями Niagara и Rock.

Процессор Niagara и системы на его основе должны быть представлены в конце 2005 или начале 2006 г. На практике он выглядит так: 90-нм технология, восемь полноценных самостоятельных ядер на кристалле, каждое ядро способно исполнять до четырех вычислительных потоков одновременно. Итак, в одном кристалле пользователям доступно 32 «виртуальных» процессора. Высокая степень интеграции позволила разместить на кристалле контроллер памяти DDR2 и четыре контроллера Gigabit Ethernet. По предварительным данным, первые образцы будут функционировать на частоте 1,2 ГГц при тепловыделении на уровне примерно 50 Вт. В дальнейшем предполагается значительно поднять тактового частоту процессора. Проект Niagara — принципиально однопроцессорное решение. В значительной степени это связано с желанием Sun отработать на этом этапе заложенные в основу Niagara принципы организации многопоточности, не отвлекая усилия на организацию еще и межпроцессорных взаимодействий. В следующем поколении, известном под условным именем Niagara 2, предполагается реализовать необходимые для построения SMP-систем компоненты. Niagara в некотором роде специализированный процессор — он разрабатывался с учетом результатов исследований об особенностях работы многопоточных приложений и наилучшим образом подходит для построения серверных систем, ориентированных на предоставление сетевых услуг. Отметим, что, по информации, предоставленной Sun, процессор Niagara будет ориентирован на использование в массовых системах нижнего ценового уровня, цена которых будет сравнима или даже ниже аналогичных по функциональности систем, построенных на базе платформы х86.

Микропроцессор Rock, появление которого намечено на 2008 г., представляет собой дальнейшее развитие идей, реализованных в проекте Niagara, однако в этой реализации «центр тяжести» перенесен на построение серверных систем, занятых интенсивной обработкой данных, для чего в архитектуру внесены соответствующие изменения.

**Список литературы**

Журнал Upgrade4\_08\_05