

Эльбрус-8С – первый отечественный восьмиядерный микропроцессор, изготовленный по технологии 28 нм

А.С. Кожин, Д.М. Альфонсо, Р.В. Деменко, Е.С. Кожин, Р.Е. Колычев, В.О. Костенко,
Н.Ю. Поляков, Е.В. Смирнова, Д.А. Смирнов, П.А. Смольянов, В.В. Тихорский

АО МЦСТ, Alexey.S.Kozhin@mcst.ru

Аннотация — Статья посвящена микропроцессору Эльбрус-8С – первому российскому восьмиядерному микропроцессору, спроектированному по технологии 28 нм. Приводятся его характеристики и структурная схема, рассматриваются основные архитектурные и микроархитектурные решения, позволившие получить высокий уровень быстродействия микропроцессора и вычислительных комплексов на его основе с учетом технологических ограничений на занимаемую площадь и рассеиваемую мощность.

Ключевые слова — Эльбрус, процессорное ядро, многоядерность, кэш-память, подсистема памяти, коммутация соединений, когерентность, ccNUMA.

I. ВВЕДЕНИЕ

В 2011 году компания МЦСТ выпустила спроектированный на базе технологии 90 нм микропроцессор Эльбрус-2С+ с тактовой частотой 500 МГц и производительностью 28 GFLOPS на операциях с одинарной точностью (FP32) – первый многоядерный микропроцессор архитектуры Эльбрус, в состав которого вошли два универсальных ядра и 4 ядра DSP [1]. Следующим проектом компании стал четырехъядерный микропроцессор Эльбрус-4С с рабочей частотой 800 МГц и производительностью более 50 GFLOPS (FP32), изготовленный по технологическим нормам 65 нм и прошедший государственные испытания весной 2014 года.

Разработка микропроцессора Эльбрус-8С была начата в конце 2011 года [2]. Ее целью стало достижение быстродействия не менее 150 GFLOPS (FP32), что в три раза превышает вычислительную мощность микропроцессора предыдущего поколения. Так как на производительность микропроцессора влияет множество его компонентов и составных блоков, некоторые из которых могут стать узким местом системы и существенно ограничить быстродействие, на этапе логического проектирования необходимо было решить ряд задач:

- увеличить производительность процессорного ядра;
- увеличить число ядер и разработать новую схему коммутации соединений;
- реализовать новую иерархию кэш-памяти;
- повысить пропускную способность подсистемы памяти и подсистемы ввода-вывода;

- оптимизировать протоколы межъядерной и межпроцессорной когерентности;
- реализовать технологии энергосбережения, отказоустойчивости и повышения выхода годных микропроцессоров.

С учетом технологических ограничений на занимаемую площадь и рассеиваемую мощность эти задачи потребовали комплексного и глубокого изменения дизайна взятого за основу микропроцессора Эльбрус-4С. В статье приводятся характеристики и особенности микропроцессора пятого поколения Эльбрус-8С, а также рассматриваются основные решения, примененные при его разработке.

II. ХАРАКТЕРИСТИКИ И СТРУКТУРНАЯ СХЕМА

Микропроцессор Эльбрус-8С спроектирован и изготовлен по технологии 28 нм. Его кристалл имеет площадь 321 мм² и содержит 2.73 млрд. транзисторов, топологический план приведен на рис. 1.

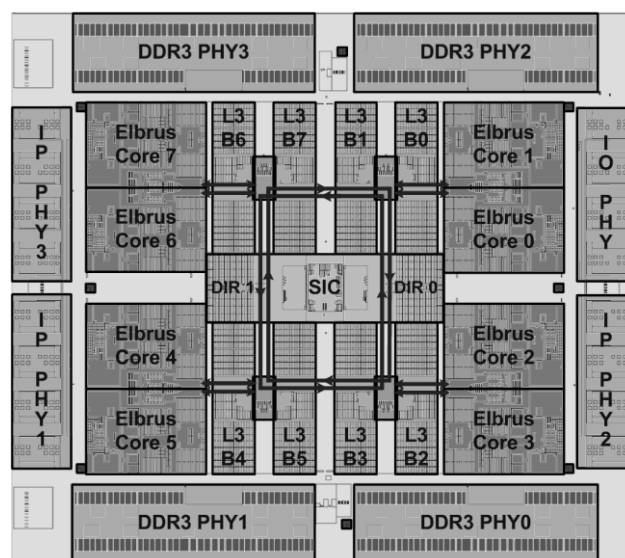


Рис. 1. Топологический план микропроцессора Эльбрус-8С (DDR3 PHY0-3 – блоки физического уровня памяти, IP PHY1,2,3 – блоки физического уровня каналов межпроцессорного обмена, IO PHY – блок физического уровня канала ввода-вывода)

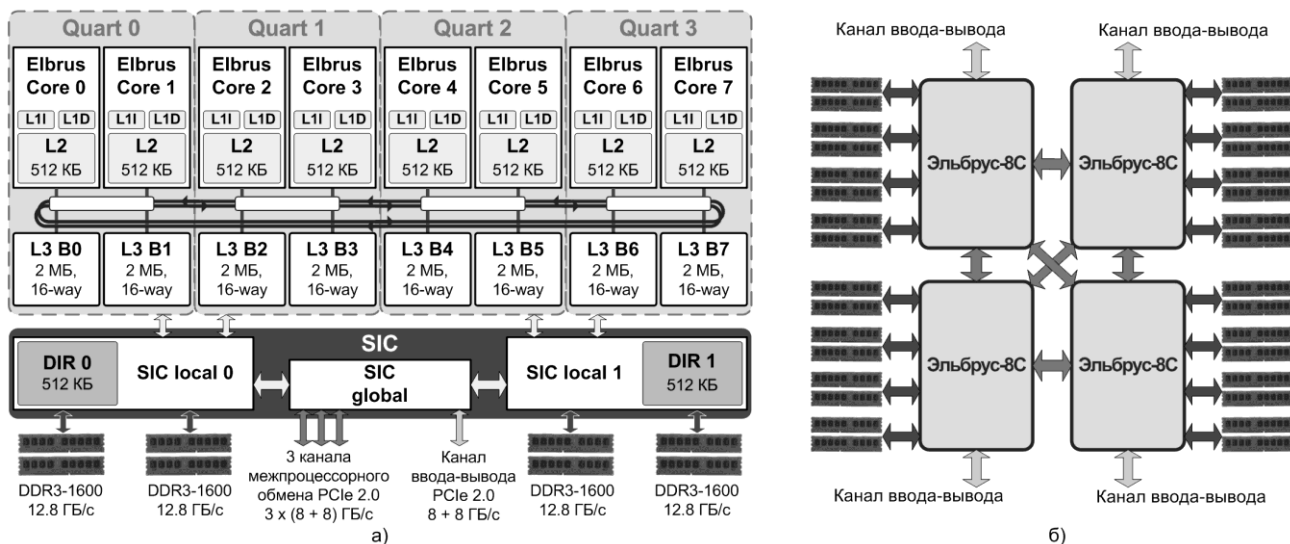


Рис. 2. а) Структурная схема микропроцессора Эльбрус-8С; б) ccNUMA система с общей когерентной памятью из четырех процессоров Эльбрус-8С

На одном кристалле размещены восемь процессорных ядер архитектуры Эльбрус (Elbrus Core 0-7), каждое из которых способно исполнять до 25 операций за такт в скалярном режиме (32-х и 64-х разрядные данные) и до 41 операции за такт в векторном режиме (упакованные 32-х разрядные данные). Ядра подключены к общей кэш-памяти третьего уровня (L3 кэш, 16 МБ), которая имеет распределенную структуру с разбиением на восемь независимых банков (L3 B0-7), причем каждое ядро имеет доступ ко всем банкам. Ядра и банки общей кэш-памяти объединены в узлы Quart (кварты) по два ядра и два банка в каждом. Кварты соединены двунаправленным буферизующим кольцом. На рис. 2а показана структурная схема микропроцессора.

Контроллер системных обменов SIC (System Interface Controller) выполняет функции доступа к памяти (SIC local 0,1) и аппаратуре ввода-вывода (SIC global), он включает в свой состав все контроллеры высокоскоростных интерфейсов процессора:

- четыре контроллера каналов оперативной памяти DDR3-1600 SDRAM с максимальной пропускной способностью 51.2 ГБ/с, которые обеспечивают установку до 64 ГБ памяти на процессор;
- три контроллера дуплексных каналов межпроцессорного обмена с пропускной способностью 8 ГБ/с в каждую сторону, позволяющие объединять до 4 микропроцессоров в одну систему с когерентной общей памятью (см. рис. 2б, возможна установка до 256 ГБ памяти на систему);
- контроллер дуплексного канала ввода-вывода с пропускной способностью 8 ГБ/с в каждую сторону, совместимый с КПИ-2 (Контроллером Периферийных Интерфейсов второй версии) собственной разработки.

Также в состав процессора входят два справочника для фильтрации запросов и снижения когерентного трафика: локальный справочник для поддержки межъядерной когерентности (включен в кэш-память третьего уровня) и глобальный справочник DIR 0,1 для поддержки межпроцессорной когерентности.

Расчетная тактовая частота разработанного микропроцессора достигает 1.3 ГГц, при этом пиковая производительность составляет 250 GFLOPS на операциях с одинарной точностью (FP32) и 125 GFLOPS на операциях с двойной точностью (FP64).

III. ПРОЦЕССОРНОЕ ЯДРО

При разработке ядра Эльбрус-8С за основу было взято ядро микропроцессора предыдущего поколения Эльбрус-4С. Для повышения производительности ядра было решено увеличить тактовую частоту и число одновременно исполняемых арифметических операций с плавающей точкой.

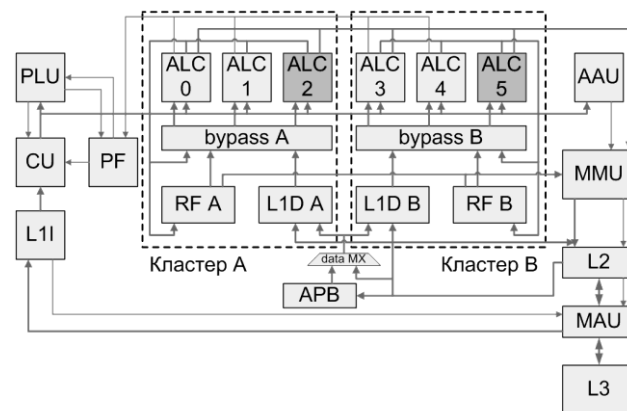


Рис. 3. Структурная схема процессорного ядра Эльбрус-8С и иерархии кэш-памяти

На рис. 3 приведена структурная схема процессорного ядра Эльбрус-8С и его отличие от ядра

микропроцессора предыдущего поколения. Ядро микропроцессора Эльбрус-4С имеет двухкластерную организацию. Каждый кластер содержит три арифметико-логических канала (ALC), локальный блок регистрового файла (RF), локальный блок кэша данных первого уровня (L1D) и блок подготовки операндов и сбора результатов (bypass). Каналы ALC имеют отдельные устройства для выполнения целочисленных операций и операций с плавающей точкой. Целочисленные операции выполняются во всех шести каналах двух кластеров. Скалярные обращения в память по считыванию выполняют ALC 0, 2, 3 и 5, по записи – ALC 2 и 5. Операции с плавающей точкой типа сложения и умножения, в том числе трехоперандные, выполняют ALC 0, 1, 3 и 4, операцию деления – только ALC 5.

Увеличение числа одновременно исполняемых арифметических операций с плавающей точкой было достигнуто за счет добавления в ALC 2 и 5 возможности исполнения операций сложения и умножения с плавающей точкой. При этом потребовалось модифицировать структуру широкой команды, а именно – организовать упаковку в нее новых слогов для кодирования новых команд ALC 2 и ALC 5 при сохранении обратной совместимости.

В результате применения обоих методов удалось повысить производительность ядра почти в 2,5 раза, при этом увеличение его площади составило всего 3,2%.

IV. ИЕРАРХИЯ КЭШ-ПАМЯТИ

Кэш-память существенно влияет на производительность микропроцессора, позволяя в максимально возможной степени сократить задержку доступа к инструкциям и данным программы. Нарращивание количества процессорных ядер только усилило роль кэш-памяти, возложив на нее дополнительную функцию обеспечения межъядерного взаимодействия.

Частная кэш-память ядра была унаследована от процессора Эльбрус-4С с некоторыми изменениями (рис. 3). Как и в предыдущем поколении, используется отдельная кэш-память первого уровня для команд

(L1I кэш, 128 КБ) и для данных (L1D кэш, 64 КБ). Кэш команд состоит из кэш-строк размером 256 байт и имеет ассоциативность 4. Размер кэш-строки виртуально индексируемого кэша данных – 32 байта, ассоциативность 4. Кэш-память второго уровня (L2 кэш) хранит и код, и данные. Ввиду технологических ограничений на площадь кристалла и новых особенностей иерархии размер L2 кэша был уменьшен до 512 КБ, в остальном он не изменился по сравнению с процессором предыдущего поколения. Размер кэш-строки L2 кэша – 64 байта, ассоциативность 4, он является неинклюзивным относительно кэш-памяти первого уровня. Устройство MAU (Memo Access Unit), подключаемое к выходу L2 кэша и управляющее обменами ядра с остальной системой, было разработано заново с целью повышения рабочей частоты и масштабируемости, а также оптимизации взаимодействия ядра с L3 кэшем.

Главной особенностью иерархии кэш-памяти нового микропроцессора стало появление общей кэш-памяти третьего уровня (L3 кэш) объемом 16 МБ [3]. Она имеет распределенную структуру и адресное разбиение (по 8, 7, 6 битам адреса) на восемь независимых банков по 2 МБ, причем любое ядро имеет доступ ко всему объему каждого банка. Размер кэш-строки данных – 64 байта, ассоциативность каждого банка 16. Разработанная схема с отдельными конвейерами тэгов и данных позволила значительно снизить динамическую мощность L3 кэша при такой высокой ассоциативности. Выбор строки-жертвы для вытеснения производится по алгоритму NRU (Not Recently Used). Память тэгов реализована на статической двухпортовой памяти с произвольным доступом, имеющей независимые порты чтения и записи. Память данных каждого банка L3 кэша разделена на два однопортовых массива шириной 32 байта, соответствующих старшей и младшей половине кэш-строки. В каждом такте можно выполнять чтение или запись только одного массива данных или одновременное чтение одного массива и запись другого. Таким образом, максимальная пропускная способность L3 кэша составляет 333 ГБ/с как по чтению, так и по записи.

Таблица 1

Время доступа в память (в наносекундах)

Процессор	Технология, частота, количество ядер	Расположение запрашиваемых данных (в скобках указан объем кэш-памяти)					L3	RAM
		Core 0		Core i, i ≠ 0				
		L1	L2	L1	L2			
Эльбрус-8С	28нм, 1.3 ГГц, 8 ядер	2.3 (64 КБ)	6.9 (512 КБ)	33.8-41.5		19.2-26.9 (16 МБ)	~ 80	
Эльбрус-4С	65нм, 0.8 ГГц, 4 ядра	3.8 (64 КБ)	11.3 (2 МБ)	57.5-61.3		n/a	~ 80	
AMD Bulldozer (Opteron 6274) [4]	32нм, 2.2 ГГц, 2x8 ядер	1.8 (16 КБ)	9.1 (2 МБ)	19.5 (in-module) 88.6 (out-module)		27.3 (8 МБ)	~ 82	
Intel SB-EP (Xeon E5-2670) [4]	32нм, 2.6 ГГц, 8 ядер	1.5 (32 КБ)	4.6 (256 КБ)	40.4	38.1	15 (20 МБ)	~ 87	

В табл. 1 приведено минимальное время доступа в память в зависимости от расположения запрашиваемых данных для разработанного микропроцессора Эльбрус-8С и его сравнение с микропроцессором предыдущего поколения Эльбрус-4С и микропроцессорами AMD Bulldozer (Opteron 6274) и Intel Sandy Bridge-EP (Xeon E5-2670), изготовленными на технологическом процессе того же уровня.

V. ПОДСИСТЕМЫ ПАМЯТИ И ВВОДА-ВЫВОДА

Подсистема памяти обеспечивает когерентный доступ процессорных ядер и внешних устройств к оперативной памяти. Как и кэш-память, она играет важную роль в достижении требуемой производительности, особенно для Memory-Bound программ. По сравнению с микропроцессором предыдущего поколения в Эльбрус-8С удвоилось число процессорных ядер и, соответственно, возросла нагрузка на память. Чтобы подсистема памяти не стала узким местом микропроцессора, количество каналов памяти DDR3-1600 было увеличено до четырех, что обеспечивает максимальную пропускную способность 51.2 ГБ/с и установку до 64 ГБ памяти на процессор. В состав микропроцессора входят три контроллера дуплексных каналов межпроцессорного обмена с пропускной способностью 8 ГБ/с в каждую сторону. Посредством этих каналов, реализованных с использованием блоков физического уровня PCI Express 2.0, возможно объединение до четырех процессоров в сeNUMA систему с общей когерентной памятью (рис. 2а).

Функции доступа к памяти выполняет локальный контроллер системных обменов SIC local. Кроме того, он является глобальной точкой сериализации запросов в многопроцессорной системе с общей памятью. В микропроцессоре Эльбрус-4С точка сериализации была способна обрабатывать не более одного запроса за процессорный такт. В новом микропроцессоре ее

пропускная способность увеличена вдвое за счет адресного разделения на два независимых кластера SIC local 0,1 (рис. 4а). Каждый кластер включает в себя два контроллера оперативной памяти MC, системный коммутатор SC, реализующий сериализацию с темпом один запрос за такт, глобальный справочник DIR и различные коммутаторы (не показаны на рисунке). Банки кэш-памяти третьего уровня имеют фиксированную привязку к кластерам: каждая пара контроллеров памяти обслуживает ближайшие четыре банка L3 кэша. Контроллеры каналов межпроцессорного обмена (IPCC 0,1,2) размещены в блоке SIC global и через коммутатор пакетов подключены к обоим кластерам SIC local.

Доступ ядер к внешним устройствам и конфигурационным регистрам системы, а также доступ внешних устройств к оперативной памяти системы посредством DMA обеспечивается через контроллер SIC global. Для связи процессора с КПИ-2 (Контроллер Периферийных Интерфейсов второй версии) используется блок физического уровня PCI Express 2.0, пропускная способность которого выросла в 4 раза по сравнению с микропроцессором предыдущего поколения Эльбрус-4С. В состав подсистемы ввода-вывода входят контроллер канала ввода-вывода (WLCC), блок конфигурационных регистров подсистемы памяти, контроллер прерываний и новый хост-контроллер (HC), разработанный для микропроцессора Эльбрус-8С. Хост-контроллер обеспечивает доступ ядер к конфигурационным регистрам и внешним устройствам, пересылку прерываний от внешних устройств и других процессоров в контроллер прерываний и доступ внешних устройств в оперативную память. Он также поддерживает упорядоченное согласно стандарту PCI Express 2.0 исполнение DMA-обращений от внешних устройств в оперативную память без потери скорости обмена [5].

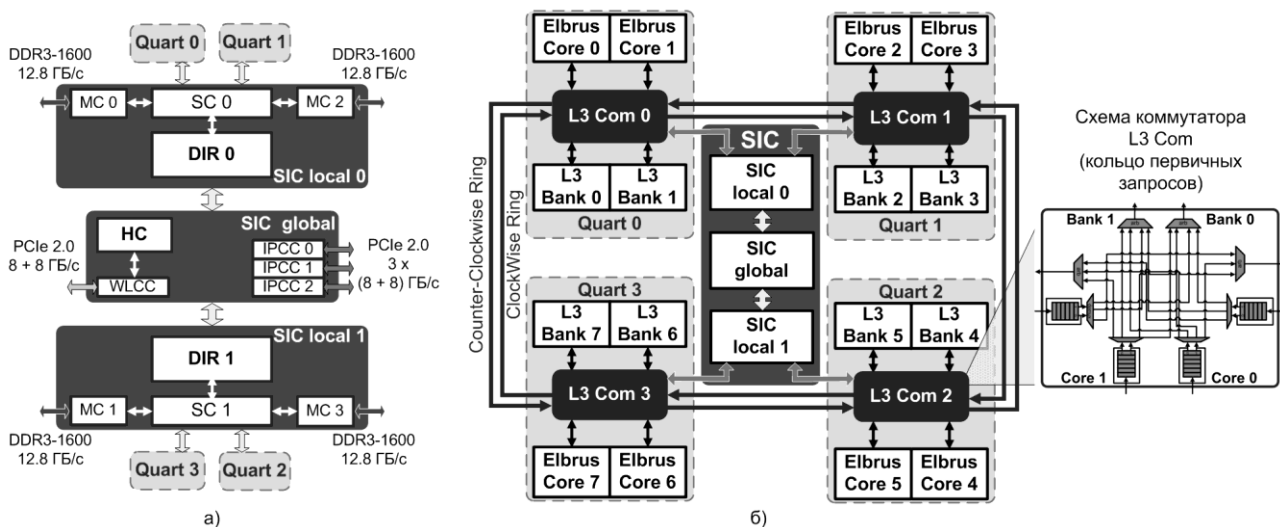


Рис. 4. а) Структурная схема контроллера системных обменов SIC; б) схема коммутации соединений микропроцессора Эльбрус-8С

VI. КОММУТАЦИЯ СОЕДИНЕНИЙ

Схема коммутации микропроцессора разделена на два уровня (рис. 4б): ядра – L3 кэш и L3 кэш – SIC. На первом уровне реализована связь каждого процессорного ядра с семью остальными и с восемью банками общей кэш-памяти третьего уровня. Разработанная схема имеет распределенную структуру и состоит из четырех узлов Quart (кварт) [6]. В кварту входят два ядра и два банка общего кэша, подключенные к централизованному коммутатору L3 Com. Кварты объединены в двунаправленное буферизирующее кольцо с передачей пакетов по часовой стрелке (ClockWise Ring) и против часовой стрелки (Counter-Clockwise Ring).

Пять независимых двунаправленных колец передачи информации, соответствующих первичным запросам, снуп-запросам, снуп-ответам, пакетам данных, подтверждениям завершений чтений, используют различные группы сигналов двунаправленной кольцевой шины. Пропускная способность кольца данных составляет 32 байта/такт в каждом направлении. Пакеты данных могут иметь формат 32 байта (половина кэш-строки) и 64 байта (целая кэш-строка), причем во втором случае пакет передается двумя неразрывными посылками по 32 байта. В пакете (независимо от его типа) указывается идентификатор абонента кольца (адрес назначения), которому он посылается. В коммутаторе L3 Com в результате анализа адреса назначения происходит дальнейшая маршрутизация пакета. Передача пакетов между абонентами кварты может выполняться независимо от их транзитного прохождения по кольцу. Для уменьшения числа конфликтов в кольце и времени доступа в L3 кэш разработано несколько механизмов оптимизации:

- Out-of-Order – возможность обгона для пакетов с разными адресами назначения;
- широковещательные пакеты – передача одинаковых снуп-запросов и пакетов с данными нескольким адресатам за одну посылку;
- упреждающая выдача заголовка данных (совместно с байпассированием данных в ядре уменьшает время попадания в L3 кэш на 20%).

Пиковая пропускная способность по чтению из общей кэш-памяти, которую может обеспечить разработанное кольцо при отсутствии конфликтов пакетов или обращении ядер в ближайшие банки, составляет 333 ГБ/с.

На втором уровне реализована связь банков L3 кэша с SIC. Каждая кварта подключена к соответствующей паре контроллеров памяти, входящих в состав одного из устройств SIC local. Коммутация происходит в два этапа: между двумя банками в кварте и между двумя квартами в SIC local. При доступе к подсистеме ввода-вывода и каналам межпроцессорного обмена коммутация между всеми четырьмя квартами осуществляется в устройстве SIC global. Суммарная пиковая пропускная способность

связей между четырьмя квартами и контроллером SIC составляет по 166 ГБ/с в каждом направлении.

VII. КОГЕРЕНТНОСТЬ

Коммутационные соединения, описанные выше, обеспечивают достаточно высокую пропускную способность для обменов как внутри чипа, так и между чипами. Тем не менее при использовании протокола когерентности на основе широковещательного снупирования они бы оказались перегружены когерентным трафиком в максимальной 32-ядерной ccNUMA системе. Для уменьшения трафика и времени доступа в память был спроектирован иерархический MOSI протокол с двумя точками сериализации запросов и использованием когерентных справочников.

В качестве точки сериализации запросов от процессорных ядер выступает общая кэш-память третьего уровня, на базе которой реализован локальный справочник с MOSI-протоколом когерентности, существенно снижающий межъядерный трафик когерентных сообщений. Принцип работы справочника основан на свойстве инклюзивности L3 кэша относительно L1 и L2 кэшей каждого ядра своего процессора. Строки, находящиеся в L1 и L2 кэшах всех ядер, обязательно присутствуют в кэш-памяти третьего уровня (возможно без актуальных данных); вытеснение строки из L3 кэша вызывает вытеснение этой строки из всех кэшей верхних уровней. Каждая кэш-строка L3 кэша дополнена информацией о ее состоянии в кэш-памяти всех процессорных ядер – вектором возможных ядер-совладельцев и указателем на владельца модифицированной строки. Эта информация вместе с памятью тэгов и состояний L3 кэша является локальным справочником. Рассылка снуп-запросов выполняется только ядрам, указанным в справочнике, а не широковещательно. Пропускная способность локального справочника составляет 8 запросов/такт. При этом адресные блокировки происходят только для запросов по записи, последовательные запросы по чтению от разных ядер для ускорения работы с общими данными могут выполняться одновременно. Доступ к контроллеру системных обменов SIC происходит только при отсутствии данных в L3 кэше и предыдущих невыполненных кэшируемых обращениях к той же кэш-строке. Такая схема позволяет значительно уменьшить нагрузку на подсистему памяти, имеющую меньшую пропускную способность.

На глобальном уровне запросы от всех процессоров сериализуются в SIC Home-процессора. Как и в процессоре предыдущего поколения, в микропроцессоре Эльбрус-8С используется глобальный справочник DIR 0,1 (рис. 2а) для уменьшения когерентного трафика и времени обработки запросов в память, но, что принципиально, не полный по памяти, а частичный [7]. Полные справочники обычно располагаются в выделенных областях оперативной памяти и содержат информацию о когерентном состоянии каждой строки памяти, тогда

как частичные справочники хранят лишь часть информации в специально отведённой для этих целей кэш-памяти. Основными доводами против варианта полного справочника послужили достаточно сильно возрастающая в этом случае сложность адресной арифметики, снижение производительности и избыточность, так как размер полного справочника пропорционален объёму всей оперативной памяти, а необходимо покрыть только суммарный объём кэш-памяти четырех процессоров.

Для увеличения эффективности покрытия глобального справочника было решено:

- не размещать в справочнике информацию о собственном процессоре, хранить информацию обо всех кэш-строках, принадлежащих памяти данного процессора и находящихся в данный момент в других процессорах;
- для каждого адресного тэга хранить информацию о двух соседних кэш-строках.

Глобальный справочник выполнен на двухпортовой SRAM-памяти, имеющей отдельные порты чтения и записи. В процессоре Эльбрус-8С находятся два экземпляра справочника: по одному на каждый SIC local. Объём каждого из них составляет 512 Кбайт (2^{17} блоков), организованных в виде кэш-памяти с ассоциативностью 16. В глобальном справочнике реализован MOSI-протокол когерентности, справочник каждого процессора покрывает 32 Мбайт памяти, то есть может полностью покрыть L3 кэши двух процессоров.

VIII. РАССЕЙВАЕМАЯ МОЩНОСТЬ, ОТКАЗОУСТОЙЧИВОСТЬ И ТЕХНОЛОГИЧЕСКИЕ ДЕФЕКТЫ

В связи с увеличением площади кристалла и переходом на новый технологический процесс 28 нм потребовалось решить ряд проблем, не связанных напрямую с достижением высокой производительности. В первую очередь стоит отметить проблему рассеиваемой мощности. Требование по энергосбережению учитывалось на всех этапах проектирования микропроцессора Эльбрус-8С, начиная с принятия микроархитектурных решений и заканчивая использованием малопотребляющих элементов и заказных блоков памяти собственной разработки [8, 9, 10]. Также для снижения рассеиваемой мощности при простое одного или нескольких процессорных ядер разработан механизм их динамического отключения под управлением операционной системы. Измерение температуры кристалла осуществляется посредством восьми термодатчиков, размещенных в разных точках кристалла.

Для защиты от единичных сбоях в кэш-памяти ядер реализованы проверка четности в небольших блоках памяти и корректирующие коды Хэмминга SECDED (Single Error Correction, Double Error Detection) в крупных блоках данных. Кроме того, код Хэмминга SECDED защищает общую кэш-память третьего

уровня, глобальный справочник и оперативную память.

Также проблемами стали полное покрытие технологических дефектов в SRAM-памяти и увеличение выхода годных чипов. Разработанный новый алгоритм системы автоматического тестирования (BIST, Built-In Self-Test) [11] обеспечивает достаточные условия для нахождения всех статистически важных типов дефектов, свойственных производству процессоров по технологическим нормам 28 нм. Неисправные элементы памяти, обнаруженные BIST, заменяются резервными элементами с помощью системы автоматической компенсации неисправностей (BISR, Built-In Self-Repair). При невозможности замены дефектный блок помечается специальной меткой как запрещённый к использованию и не влияет на корректность работы процессора, но при этом несколько снижается производительность. Реализованные методики резервирования и маскирования позволяют использовать микропроцессоры с дефектами в SRAM-памяти и таким образом увеличивают выход годных процессоров.

IX. ЗАКЛЮЧЕНИЕ

Цель работы по проектированию универсального микропроцессора с производительностью не ниже 150 GFLOPS на операциях с одинарной точностью была достигнута за счет создания сбалансированного дизайна микропроцессора и решения целого ряда задач на архитектурном и микроархитектурном уровнях, а также преодоления проблем, характерных для нового технологического процесса 28 нм. Пиковая производительность разработанного микропроцессора Эльбрус-8С составляет 250 GFLOPS на операциях с одинарной точностью при расчетной тактовой частоте 1.3 ГГц. Первые инженерные образцы были получены в ноябре 2014 года, начало серийного производства микропроцессора намечено на 2016 год.

БЛАГОДАРНОСТИ

Авторы статьи выражают благодарность всем сотрудникам АО МЦСТ и ПАО «ИНЭУМ им. И.С. Брука», принимавшим участие в проектировании, верификации и отладке микропроцессора Эльбрус-8С.

ЛИТЕРАТУРА

- [1] Исаев М.В., Кожин А.С., Костенко В.О., Поляков Н.Ю., Сахин Ю.Х. Двухъядерная гетерогенная система на кристалле «Эльбрус-2С+» // Вопросы радиоэлектроники / Вып. 3. Сер. ЭВТ. М., 2012. С. 42-52.
- [2] Kostenko V.O., Kozhin A.S., Polyakov N.Y., Slesarev M.V., Tikhorskiy V.V., Sakhin Y.K. Elbrus-8C: The Latest Yield from MCST and MIPT Collaboration // 2015 International Conference on Engineering and Telecommunication (EnT). 2015. P. 67-68.
- [3] Кожин А.С., Кожин Е.С., Костенко В.О., Лавров А.В. Кэш третьего уровня и поддержка когерентности микропроцессора «Эльбрус-4С+» // Вопросы

- радиоэлектроники / Вып. 3. Сер. ЭВТ. М., 2013. С. 26-38.
- [4] Molka D., Hackenberg D., Schöne R. Main memory and cache performance of Intel Sandy Bridge and AMD Bulldozer // Proceedings of the workshop on Memory Systems Performance and Correctness, MSPC '14. 2014. Article No. 4.
- [5] Перов Д.Ю., Поляков Н.Ю. Обеспечение упорядоченности выполнения DMA-операций в NUMA-системах методом предварительного кэширования // Вопросы радиоэлектроники / Вып. 3. Сер. ЭВТ. М., 2013. С. 38-47.
- [6] Кожин А.С., Сахин Ю.Х. Коммутация соединений процессорных ядер с общим кэшем третьего уровня микропроцессора «Эльбрус-4С+» // Вопросы радиоэлектроники / Вып. 3. Сер. ЭВТ. М., 2013. С. 5-14.
- [7] Варакин В.Н., Исаев М.В., Сахин Ю.Х. Оптимизация межпроцессорного протокола когерентности с помощью справочника микропроцессора «Эльбрус-4С+» // Вопросы радиоэлектроники / Вып. 3. Сер. ЭВТ. М., 2013. С. 14-26.
- [8] Мороз Я.Н. Основные этапы и методы физического проектирования микропроцессора «Эльбрус-8С» // Вопросы радиоэлектроники / Вып. 3. Сер. ЭВТ. М., 2016.
- [9] Потовин Ю.М., Соин С.А. Разработка быстродействующего блока памяти с ассоциативной выборкой // Проблемы разработки перспективных микро- и нанoeлектронных систем - 2014. Сборник трудов / под общ. ред. академика РАН А.Л. Стемпковского. М.: ИППМ РАН, 2014. Часть IV. С. 29-32.
- [10] Бычков И.Н. Разработка корпуса многоядерного процессора на основе керамической коммутационной платы // Проблемы разработки перспективных микро- и нанoeлектронных систем - 2014. Сборник трудов / под общ. ред. академика РАН А.Л. Стемпковского. М.: ИППМ РАН, 2014. Часть II. С. 163-166.
- [11] Альфонсо Д.М., Исаев М.В., Костенко В.О. Разработка системы тестирования и повышения выхода годной продукции для кэш-памяти микропроцессора // Вопросы радиоэлектроники / Вып. 3. Сер. ЭВТ. М., 2014. С. 106-118.

Elbrus-8C: the first Russian 28 nm 8-core processor

A.S. Kozhin, D.M. Alfonso, R.V. Demenko, E.S. Kozhin, R.E. Kolychev, V.O. Kostenko, N.Yu. Polyakov, E.V. Smirnova, D.A. Smirnov, P.A. Smolyanov, V.V. Tikhorskiy

АО MCST, Alexey.S.Kozhin@mcst.ru

Keywords — Elbrus, core, multicore, cache hierarchy, memory subsystem, on-chip interconnect, coherency, ccNUMA.

ABSTRACT

This paper is concerned with the first Russian 28 nm 8-core processor Elbrus-8C developed by MCST company and released in 2015.

The goal of Elbrus-8C processor was to achieve more than 150 GFLOPS single-precision floating-point performance, making it around 3 times faster than its predecessor quad-core Elbrus-4C processor. Because microprocessor performance depends on many of its subblocks, any one of them can become a performance bottleneck. In order to meet the requirements, the MCST's team had to solve the following design challenges:

- 1) improve floating-point performance of each processor core,
- 2) integrate 8 cores on a single chip doubling this number from previous generations,
- 3) design high-bandwidth, low-latency scalable on-chip interconnect,
- 4) augment cache hierarchy with new level of cache,
- 5) increase memory and I/O subsystems bandwidth to feed the processor cores,
- 6) develop an efficient cache coherence protocol for intra- and inter-processor communication,

- 7) improve energy efficiency of a chip,
- 8) increase reliability and die yield characteristics.

All these design challenges along with technological constraints on the processor area and power required complex and extensive changes of Elbrus-4C processor design taken as a reference point.

MCST's flagship microprocessor Elbrus-8C is an ASIC design implemented at 28 nm technology node. The 321 mm² die contains 2.73 billion transistors. The target microprocessor's clock frequency is 1.3 GHz. In this case the peak performance is 250 GFLOPS for single-precision (FP32) and 125 GFLOPS for double-precision (FP64).

In this paper, the microarchitecture, floorplan and key features of Elbrus-8C are introduced. The main technical decisions and technologies, which enable to achieve high performance with area and power constraints, are discussed.

REFERENCES

- [1] Isaev M.V., Kozhin A.S., Kostenko V.O., Polyakov N.Yu., Sahin Yu.Kh. Dual-core heterogeneous system-on-chip «Elbrus-2S+» // Voprosy radiojelektroniki / No. 3, ser. EVT, Moscow, 2012, pp. 42-52 (in Russian).
- [2] Kostenko V.O., Kozhin A.S., Polyakov N.Yu., Slesarev M.V., Tikhorskiy V.V., Sahin Yu.Kh. Elbrus8C: The Latest Yield from MCST and MIPT Collaboration // 2015 International Conference on Engineering and Telecommunication (EnT). 2015. P. 67-68.

- [3] Kozhin A.S., Kozhin E.S., Kostenko V.O., Lavrov A.V. L3 cache and cache coherence support in «Elbrus-4C+» microprocessor // *Voprosy radioelektroniki* / No. 3, ser. EVT, Moscow, 2013, pp. 26-38 (in Russian).
- [4] Molka D., Hackenberg D., Schöne R. Main memory and cache performance of Intel Sandy Bridge and AMD Bulldozer // *Proceedings of the workshop on Memory Systems Performance and Correctness, MSPC '14*. 2014. Article No. 4.
- [5] Perov D.Yu., Polyakov N.Yu. DMA operations order support in NUMA systems by means of DMA caching method // *Voprosy radioelektroniki* / No. 3, ser. EVT, Moscow, 2013, pp. 38-47 (in Russian).
- [6] Kozhin A.S., Sahin Yu.Kh. Signal commutation between cores and shared L3 cache in «Elbrus-4C+» // *Voprosy radioelektroniki* / No. 3, ser. EVT, Moscow, 2013, pp. 5-14 (in Russian).
- [7] Varaksin V.N., Isaev M.V., Sahin Yu.Kh. Optimization of multiprocessing memory coherence protocol with directory in «Elbrus-4C+» // *Voprosy radioelektroniki* / No. 3, ser. EVT, Moscow, 2013, pp. 14-26 (in Russian).
- [8] Moroz Y.N. Osnovnie etapy i metody fizicheskogo proektirovaniya mikroprocessora «Elbrus-8C» (Main stages and methods of physical design of the microprocessor «Elbrus-8C») // *Voprosy radioelektroniki* / No. 3, ser. EVT, Moscow, 2016 (in Russian).
- [9] Potovin Yu.M., Soin S.A. High-speed content addressable memory block design // *Problems of Perspective Micro- and Nanoelectronic Systems Development*. 2014. *Proceedings* / edited by A. Stempkovsky, Moscow, IPPM RAS, 2014. Part IV. P. 29-32 (in Russian).
- [10] Bychkov I.N. Design of IC package with ceramic substrate for multicore processor // *Problems of Perspective Micro- and Nanoelectronic Systems Development - 2014*. *Proceedings* / edited by A. Stempkovsky, Moscow, IPPM RAS, 2014. Part II. P. 163-166 (in Russian).
- [11] Alfonso D.M., Isaev M.V., Kostenko V.O. Razrabotka sistemy testirovaniya i povysheniya vyhoda godnoj produkcii dlja kjesh-pamjati mikroprocessora // *Voprosy radioelektroniki* / No. 3, ser. EVT, Moscow, 2014, pp. 106-118 (in Russian).