Микросхемы Huawei для массивов All-Flash

Обзор функциональных особенностей и преимуществ использования специализированных микросхем компании Huawei для обработки протоколов в составе платы SmartIO, а также для управления в составе SSD-накопителей при реализации AFA-массивов.

Кровяков Иван – технический директор направления IT, Департамент корпоративных решений, компания Huawei.

Введение

Многих интересует вопрос, почему компания Huawei решила разрабатывать собственные микросхемы для своих систем хранения данных? Компания Avago приобрела компанию LSI, компания Cavium – компанию QLogic, а компания Broadcom приобрела компанию Brocade, чтобы получить технические возможности микросхем для определенного вида оборудования. Итак, почему же компания Huawei, вопреки ожиданиям, решила разрабатывать собственные микросхемы вместо того, чтобы просто приобрести компанию-производителя микросхем? Какие виды микросхем для СХД компания Huawei разрабатывает самостоятельно? Как эти микросхемы будут усовершенствоваться в будущем? В этой статье мы рассмотрим общие направления, которые взяла и которых намерена придерживаться компания Huawei.

Положение на рынке

Стимул для разработки собственных микросхем

Деятельность предприятий во всем мире сопровождается огромными потоками информации, а это означает, что для хранения данных требуются высокообрабатывающие, высокопроизводительные и высокоэффективные системы хранения. Технология All-Flash способствует цифровым трансформациям в силу присущих ей преимуществ в плане производительности, необходимой критическим сервисам.

С быстрым развитием технологий флеш-памяти накопители данных становятся более качественными, чем количественными (быстродействие накопителей 3D XPoint в 1000 раз превышает быстродействие накопителей NAND). Основным требованием для разработки носителей информации и приложений, таких как AR/VR, является более низкая задержка в сети. Стабильный период удваивания производительности процессоров, равный 18 месяцам, согласно закону Мура, теперь расширяется по мере замедления роста производительности процессора (увеличивается разрыв между вычислительной мощностью и возможностями обработки).

Несбалансированность разработок в сферах накопителей данных, сетевых интерфейсов и процессоров диктует необходимость изменения технологий хранения.

Поняв необходимость революции в технологиях хранения данных, компания Huawei предложила для новых решений взять All-IP в качестве основы, а All-Cloud и All-Flash – в качестве движущей силы. Это позволит оптимизированным микросхемам, сетевому оборудованию и вертикальной интеграции в приложениях сыграть дополнительную роль для оправдания ожиданий пользователей услуг. Вертикальная интеграция позволит максимально использовать ресурсы, чтобы обеспечить клиентам более высокую производительность продукта.

Инновации в технологии самостоятельно разработанных микросхем

В 2016 году компания Huawei выпустила новое поколение дисковых массивов – OceanStor Dorado V3 – на базе флеш-памяти, отличающихся производительностью до 4 миллионов IOPS при сохранении стабильной задержки в 0,5 мс. Новые массивы Dorado способны в полной мере использовать эффективность флеш-памяти, чтобы удовлетворить требования к обработке услуг для критически важных сервисов в условиях высокой нагрузки. Компания Huawei осуществила комплексную оптимизацию своих основных микросхем, включая клиентские микросхемы обработки протоколов, микросхемы ускорения обработки операций ввода-вывода и микросхемы управления SSD-накопителями, с целью ориентировать их на использование в платформах хранения данных на базе флеш-памяти. Данная оптимизация позволит на 200% повысить производительность сквозной передачи.

Микросхемы обработки протоколов

Интеграция нескольких протоколов на плате Huawei SmartIO, сокращение количества соединительных кабелей, упрощение создания сетей и снижение TCO

На платах SmartIO используются микросхемы производства компании Huawei, выполняющие обработку таких протоколов хранения, как 8G/16G FC, 10G FCoE, 10GE и iWARP. Это позволит клиентам комбинировать трафик данных IP и FC на одной интерфейсной микросхеме.

Возможности конвергенции на уровне сети обеспечивают использование 10GE для поддержки протоколов FCoE/iSCSI/

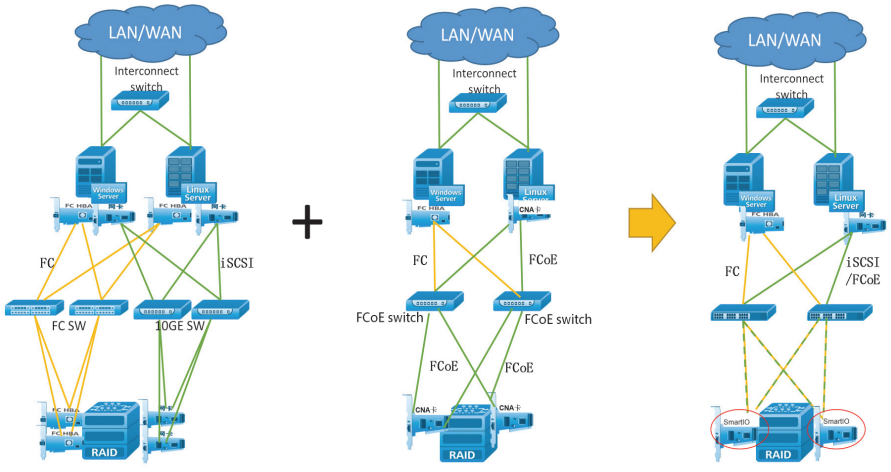

Рис. 1. Использование микросхем компании Huawei на платах SmartIO позволяет реализовать конвергенцию протоколов FCoE/iSCSI/iWARP/CIFS/NFS на базе 10GE без замены каких-либо физических компонентов.

Табл. 1. Результаты тестирования массива OceanStor Dorado V3.

I/O 8 КБ	TCP/IP более 10GE	iWARP более 10GE
Задержка	115 мкс	42 мкс

iWARP/CIFS/NFS без замены каких-либо физических компонентов. В сети стандарта FC 10GE или 8/16 ГБ потребуется замена только компонентов оптического модуля (без замены плат, с поддержкой преобразования любого протокола). При этом количество необходимых кабелей уменьшится на треть, а количество физических компонентов на интерфейсных платах – на 75%, что в свою очередь минимизирует требуемые инвестиции.

Технология RDMA, снижение задержки каналов связи на 60%

Удалённый прямой доступ к памяти (Remote Direct Memory Access, RDMA) – функция, которая позволяет компьютеру напрямую передавать данные в память другого компьютера по сети без какого-либо влияния на работу операционной системы. Это значит, что для завершения данной операции возможности обработки компьютеров не используются, в связи с чем пространство шины освобождается, и циклы процессора сокращаются, таким образом повышая производительность системы приложений. Микросхемы обработки протоколов хранения производства Huawei поддерживают на аппаратном уровне RDMA, передачу данных и коммутацию между несколькими контроллерами памяти посредством RDMA, в результате чего задержка канала сокращается на 60% и более, а также значительно повышается эффективность обработки в условиях высокой нагрузки услуг и одновременного доступа.

Результаты тестирования массива OceanStor Dorado V3 при одновременном доступе 8 станций с размером блока ввода-вывода 8 кбайт на скорости 10GE представлены в табл. 1.

Снижение перегрузки в глобальной сети (WAN) с оптимизацией параметров, направленной на увеличение пропускной способности при выполнении удаленной репликации в сложных сетевых сценариях

Оптимизация WAN направлена на уменьшение объема данных, передаваемых по таким сетям, путем применения различных технических подходов. Осуществляется оптимизация передачи данных и использования ресурсов полосы пропускания. Оптимизацию можно реализовать с одной стороны или с двух сторон (при односторонней оптимизации используется функция управления трафиком и технология оптимизации TCP, в то время как при двусторонней оптимизации используется кэш и технологии сжатия).

Микросхемы обработки протоколов хранения данных компании Huawei имеют встроенные функции контроля трафика QoS и предотвращения перегрузки TCP. В сложных сетевых сценариях (например, в случае подключения многочисленных вычислительных и коммуникационных сетей удаленно через LAN или MAN или даже через большой географический район) охватываемая область может варьироваться от нескольких десятков до нескольких сотен километров при подключении объектов связи в разных городах. Информация, собранная в режиме реального времени, о задержке времени оборота (RTT) сетевых пакетов, коэффициенте потери пакетов, ECN и других характеристиках в сочетании с различными алгоритмами предотвращения перегрузки TCP используется для настройки стратегий приема и передачи, включая повторную передачу, окно буфера приема/передачи и интеллектуальный контроль трафика. Настройка стратегий дает возможность предотвратить перегрузку определенных каналов, а также выполнить сброс вручную и динамически, в результате чего сетевые платы смогут работать быстрее, чем обычно. Ускорение работы плат повышает производительность на 65–400% в сложных сетях WAN.

Микросхема ускорения процесса обработки ввода-вывода

В эпоху облачных сервисов объемы данных растут быстрее, чем ожидалось, а также увеличивается количество дублированной информации. Дублированные данные, как правило, не имеют ценности для организаций. Они занимают большое пространство в системах хранения, увеличивают расходы на потребление электроэнергии и охлаждение, а также снижают производительность доступа. Это приводит к нерациональному использованию ограниченных ИТ-ресурсов. Организациям требуется быстрый доступ к своим данным и способы, которые позволят сократить объем дублированных данных.

Для решения этой проблемы в отрасли широко применяются технологии дедупликации и сжатия. Однако эти функции потребляют большой объем ресурсов ЦП из-за выполнения многочисленных алгоритмов определения, сжатия и восстановления характерных фрагментов данных. Использование этих функций приводит к снижению качества обслуживания. Кроме того, клиенты замечают отсутствие изменения объема зарезервированного пространства для хранения их данных, а также отсутствие экономии денежных средств относительно стартовых инвестиций. Также недостатком является тот

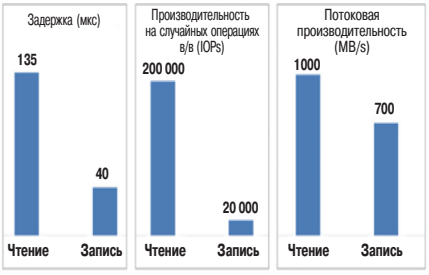

Рис. 2. Справочные данные по производительности для твердотельных накопителей производства Huawei.

факт, что срок службы SSD зависит от количества записанных данных.

Микросхемы ускорения процесса обработки ввода-вывода производства Huawei имеют интегрированные механизмы алгоритмов сжатия и восстановления данных, которые передают задачи, потребляющие большое количество вычислительных ресурсов, механизму алгоритмов, уменьшая нагрузку на процессоры. В ходе тестирования с большим количеством последовательных операций ввода-вывода, занятое на процессорах пространство памяти уменьшилось на 24,6%, IOPS увеличилось на 342,4%, а задержка сократилась на 77,4%.

Микросхемы управления твердотельными накопителями (SSD)

В SSD-накопителях компании Huawei используется новое поколение собственных микросхем управления с процессорами от компании Cortex-A9. Эти микросхемы поддерживают оперативную память DDR4 и до 18 каналов флеш-памяти NAND. Аппаратная технология FTL (Flash Translation Layer) позволяет ускорить процесс обработки ввода-вывода и обеспечивает ведущий в отрасли показатель 200 000 IOPS (рис. 2).

Аппаратная технология FTL сокращает время задержки на 20% по сравнению со средними показателями в отрасли при небольшой загрузке системы

Утилита FTL является неотъемлемой частью структуры SSD. Она предназначена для установления соответствия между

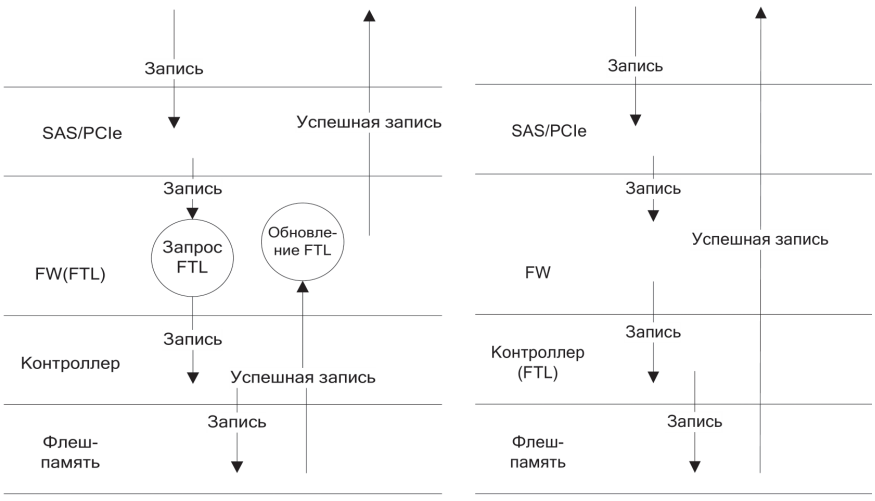

Рис. 3. Сравнение логики реализации технологии FTL программно и аппаратно.

LBA пользователя и физической страницей в SSD. Когда пользователь считывает и записывает данные, FTL передает адрес LBA. После того как SSD получает адрес, запрашивается таблица FTL для чтения данных с соответствующей физической страницы адреса LBA. Для сравнения, когда данные считываются с обычных твердотельных накопителей, встроенное программное обеспечение управления находит соответствующий физический адрес LBA, а затем считывает данные из флеш-памяти и передает их на хост. При записи данных таблица соответствия FTL обновляется после того как программное обеспечение завершит запись. SSD-накопители производства Huawei обеспечивают ускорение управления таблицами FTL на аппаратном уровне. Все операции чтения и записи в FTL выполняются аппаратным обеспечением, что сокращает количество программных операций и задержку ввода-вывода (рис. 3). При небольшой загруженности системы задержка сокращается до 40 мкс (на 20% ниже среднего показателя в отрасли).

Технология FlashLink обеспечивает задержку в 0,5 мс во всем флеш-массиве

Технология FlashLink была реализована в ходе разработок твердотельных накопителей и операционной системы компании Huawei. Она обеспечивает вертикальную оптимизацию дискового управления между аппаратным и программным обеспечением, а также стабильность задержки в 0,5 мс во всей системе хранения данных OceanStor Dorado V3 на базе флеш-памяти (рис. 4).

При классификации наиболее востребованных («горячих») данных и менее востребованных («холодных») данных информация записывается в разные разделы. Собственные операционные системы хранения данных обладают функцией определения «температуры» данных и маркировки при доступе к ним на SSD-накопителях. Микросхемы управления SSD-накопителями направляют данные по определенному пути хранения в соответствии с метками «горячие»/«холодные», тем самым уменьшая объем данных, которые необходимо перемещать при удалении, и предотвращая увеличение объема записи примерно на 40%, а задержку – на 20%.

Функция планирования приоритетов ввода-вывода предоставляет множество возможностей для обеспечения низкой и стабильной задержки. Операционная система производства Huawei определяет уровень приоритета ввода-вывода. Например, хост определяет, что приоритет запроса на чтение данных выше, чем приоритет запроса на очистку кэша. Приоритет запроса на очистку выше, чем запроса на асинхронную репликацию фоновых копий. Приоритет ввода-вывода передается на SSD вместе с запросами на чтение и запись. Далее микросхема управления SSD-накопителями обрабатывает операции ввода-вывода в соответствии с приоритетами, тем самым обеспечивая комплексное управление приоритетами ввода-вывода. Приоритетным услугам гарантируется первоочередная обработка при чтении и записи данных.

Заключение

Компания Huawei намерена и далее инвестировать средства в разработку своих собственных микросхем, которые позволят организациям взять под контроль потоки рабочей информации и добиться цифровой трансформации.

Микросхемы обработки протоколов, в которых реализована концепция «All-IP», участвуют в процессах на стороне клиента и на стороне сервера, в коммутации, репликации и сети со схемой «активный-активный», обеспечивая при этом низкую задержку всей конструкции. Микросхемы ускорения процесса обработки ввода-вывода, в которых используется подход SOC, обеспечивают функции интегрированного хранения и повышения производительности сети с постоянной оптимизацией возможностей вычисления, хранения и пропускной способности сети. Микросхемы управления, оптимизированные для SSD-накопителей, ориентированы на эволюцию для полной реализации преимуществ носителей информации нового поколения.

Кровяков Иван,
компания Huawei

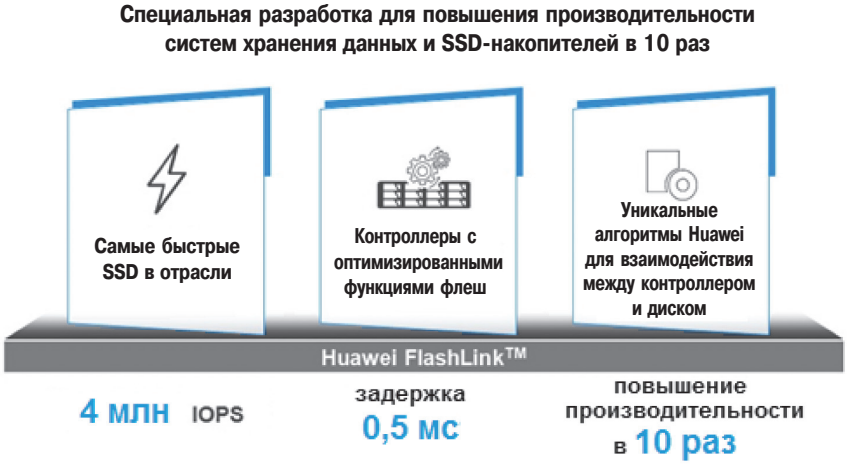

Рис. 4. Технология FlashLink обеспечивает вертикальную оптимизацию дискового управления между аппаратным и программным обеспечением.

Выпущен первый стандарт 3GPP 5G NR

Декабрь 2017 г. – На пленарном заседании Группы по разработке технических спецификаций сетей радиодоступа (TSG RAN) консорциума 3GPP в Лиссабоне были успешно представлены первые практически реализуемые спецификации «нового радио» (New Radio (NR)) для сетей 5G. В его разработке участвовали компании AT&T, BT, China Mobile, China Telecom, China Unicom, Deutsche Telekom, Ericsson, Fujitsu, **Huawei,** Intel, KT Corporation, LG Electronics, LG Uplus, MediaTek Inc., NEC Corporation, Nokia, NTT DOCOMO, Orange, Qualcomm Technologies, Inc., Samsung Electronics, SK Telecom, Sony Mobile Communications Inc., Sprint, TIM, Telefonica, Telia Company, T-Mobile USA, Verizon, Vodafone и ZTE. Это событие создает фундамент, который позволит мировой отрасли мобильной связи начать повсеместное коммерческое использование «нового радио» для сетей 5G уже в 2019 году.

27 февраля 2017 года в Барселоне лидеры мировой отрасли мобильной связи выступили за ускоренное воплощение в жизнь программы стандартизации NR для сетей 5G, что стало важным промежуточным этапом на пути к подготовке первых реализуемых спецификаций для стандарта Non-Standalone 5G NR (неавтономная архитектура 5G NR). На пленарном заседании рабочей группы 3GPP RAN в хорватском Дубровнике 9 марта был согласован сжатый график, а первые спецификации были представлены как часть серии документов 3GPP Release 15.

Эти стандарты стали важнейшей вехой в развитии 5G NR, значительно расширив возможности систем 3GPP и вертикального рынка. Консорциум 3GPP планирует продолжить разработку серии документов Release 15, в том числе включить в нее поддержку стандарта Standalone 5G NR, который также был одобрен в Дубровнике. Спецификации нижнего уровня 5G NR были разработаны для единообразной поддержки стандартов Standalone и Non-Standalone 5G NR, чтобы консорциум 3GPP мог построить крупномасштабную единую экосистему 5G NR. Мы признательны за огромные усилия, которые 3GPP приложила к выполнению этого сложного графика стандартизации.

Ян Чаобинь *(Yang Chaobin)*, президент направления 5G компании Huawei, сказал: «Как один из ключевых игроков, Huawei взяла на себя обязательство разработать единый глобальный стандарт 5G. Благодаря успешному сотрудничеству и объединению усилий со всемирными организациями, включая правительства, регулирующие органы, исследовательские организации, научные круги, отрасли и многие другие секторы, этап 1 стандартизации 3GPP 5G NR завершился с большим успехом. Huawei будет продолжать работать с глобальными партнерами, чтобы довести 5G до периода масштабного глобального коммерческого развертывания с 2018 года».