

# “КРОК”: Violin Memory уже в России



**Вадим Болотнов** – директор Центра решений “КРОК” на базе технологии EMC, менеджер по развитию решений “КРОК”.

Компания “КРОК” подписала соглашение с американской компанией Violin, в соответствии с которым получила эксклюзивный статус авторизованного реселлера и сервис-провайдера Violin Memory (с собственным складом запчастей), а также стала единственной компанией в России, предлагающей возможность проведения пилотного тестирования систем Violin Memory Array. В конце марта 2013 г. состоялись публичные тест-драйвы решений Violin в Москве.

Прежде всего, не стоит путать Violin Memory Array с флеш-карточками для серверов. Violin Memory – это законченная СХД корпоративного класса (рис. 1, 2), занимающая 3 серверных юнита, основой которой является так называемая “флеш-фабрика” (Flash fabric) – тысячи флеш-чипов памяти SLC или MLC в зависимости от модели (табл. 1). Компания Violin практически с момента своего основания заключила стратегическое партнерство с компанией Toshiba – производителем и разработчиком флеш-памяти. Это позволило построить архитектуру системы снизу вверх, на каждом уровне учитывая внутреннюю специфику флеш-памяти. Чипы организованы в PCI-E модули VIMM (Violin Intelligent Memory Module), напоминающие большую планку оперативной памяти. Таких модулей в системе может быть 20, 40 или 60. Помимо этого 4 VIMM всегда поставляются в качестве модулей горячей замены (Hot Spare). Каждый VIMM кроме флеш-чипов содержит выделенный контроллер

управления и DRAM-память для хранения и обработки служебных таблиц. Эти таблицы постоянно обновляются, и расположение их на флеш-памяти привело бы к преждевременному выходу модулей из строя. Наличие их внутри VIMM позволяет контроллеру оперативно решать задачу балансировки по всем чипам и следить за их равномерным износом. Кто-то может спросить: “Ведь DRAM – энергозависимая память, что происходит при сбое питания?”. Никаких дополнительных батарей в системе не предусмотрено, но 2 выделенных контроллера питания устроены так, что остаточного заряда электричества в блоках питания хватает на то, чтобы сохранить все данные в энергозависимой памяти 60 раз подряд.

Если взглянуть на уровень выше, VIMM объединяются по 5 модулей в так называемые vRAID-группы для защиты от выхода каждого из них из строя. За коммутацией и балансировкой данных по всей флеш-фабрике следят также 2 отдельных модуля в виде компактных x86-серверов. За логическое представление данных отвечают еще два – так называемые Memory Gateways. В предыдущей линейке Violin Memory 3000 эти модули представляли собой отдельные серверы. В этот раз разработчикам удалось найти место внутри системы и для них.

Violin Memory Array имеет несколько интерфейсов подключения. Во-первых, 4xPCI-E – для подключения напрямую к серверам, исключая все возможные задержки в канале. Кроме того, поддерживаются 8x8Гб/с Fiber Channel – для включения в большинство сетей хранения, 8x10 GbE iSCSI, а также 8x40Gbit/s 4xQDR Infiniband – для использования в инфраструктуре ЦОД. Стоит отметить, что вся внутренняя коммутация компонентов происходит по нескольким шинам PCI-E, что обеспечивает минимальную задержку и теоретическую пропускную способность 80 Гбит/с. Все компоненты Violin Memory дублируются, а единственный вращающийся элемент – это 6 вентиляторов, из которых для работы системы достаточно только трех.

Одна из ключевых особенностей массивов Violin – автоматическая балансировка нагрузки на уровне флеш-памяти – VIMM, что позволяет увеличивать на-

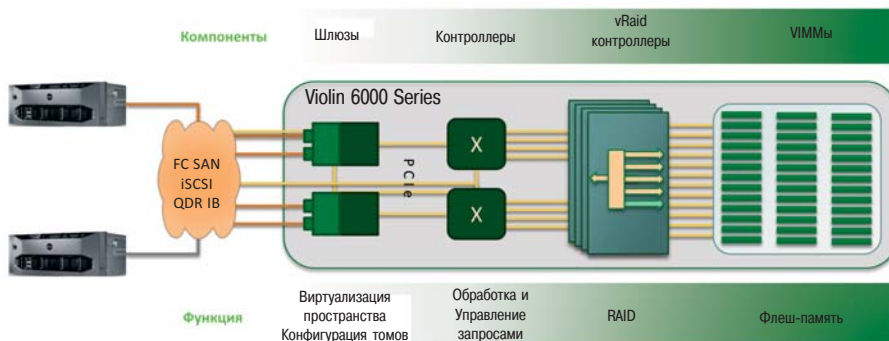
**Табл. 1.** Характеристики массивов Violin для MLC- и SLC-технологий чипов.

|  | 6212  | 6222      | 6232      | 6606                     | 6611      | 6616    |
|--|---|-----------|-----------|--------------------------|-----------|---------|
| Форм-фактор                                  | 3U  |           |           | 3U                       |           |         |
| Тип флеш-памяти                              | Объем (MLC)   |           |           | Производительность (SLC) |           |         |
| Raw емкость (GB)                             | 13.2TB  | 24.2TB    | 35.2TB    | 6.6TB                    | 12.1TB    | 17.6TB  |
| I/O подключения                              | 8Gb FC, 10GbE iSCSI, 40 Gb IB, PCIe G2  |           |           |                          |           |         |
| Максимум 4KB IOPS (смешанный режим)          | 200K IOPS   | 350K IOPS | 500K IOPS | 450K IOPS                | 800K IOPS | 1M IOPS |
| Максимальная полоса пропускания (100% Reads) | 1.5GB/s   | 2.5GB/s   | 4GB/s     | 3GB/s                    | 3.5GB/s   | 4GB/s   |
| Номинальная латентность                      | 500 мс (чтение/запись)  |           |           | 250 мс (чтение/запись)   |           |         |
| Энергопотребление                            | 1355Вт  | 1580Вт    | 1723Вт    | 1119Вт                   | 1328Вт    | 1693Вт  |
| Поддерживаемые ОС /доп. Функ-ть              | RHEL, SLES, Windows, VMware, Hyper-V, Citrix, AIX, Solaris SPARC, Solaris x86, HP-UX, OpenVMS / VAAI, vCenter Plug-In |           |           |                          |           |         |

грузку до максимального значения при практически неизменном времени отклика. Другая важная особенность в том, что алгоритмы Violin Memory построены так, что система одинаково быстро справляется как с чтением, так и записью. Это является ее ключевым преимуществом перед существующими аналогами, использующими SSD-диски, для которых даже небольшой процент активной записи (например, 10% – запись, 90% – чтение) в ряде случаев может приводить к многократному (в отдельных случаях на порядок и более) снижению производительности.

Целью тест-драйвов была проверка заявленных возможностей Violin Memory Array. *В первую очередь*, это стабильное время отклика – от 0,2 до 0,5 мс независимо от типа и степени нагрузки на систему. *Во-вторых*, важно было выяснить, как долго система будет справляться с высокой нагрузкой, не случится ли в определенный момент спада, связанного с внутренними механизмами очистки флеш-памяти. *И, в-третьих*, система, на которой заказчик располагает критичные бизнес-приложения, должна выдерживать сбой ключевых компонентов, сохраняя заявленные характеристики производительности.

Наша демо-система Violin Memory 6606 является младшей моделью в линейке на SLC-чипах. В ее составе 20 модулей VIMM по 256 Гбайт и 4 модуля горячей замены. Емкость системы после форматирования составила 2,6 Тбайт. Это наи-



**Рис. 1.** Архитектура решений Violin Memory Array.



**Рис. 2.** Внешний вид массивов Violin.

меньшая по объему СХД из всех представленных, но для проведения демонстраций этого вполне достаточно. Максимальная емкость одной системы после форматирования на текущий момент составляет 21 Тбайт с использованием памяти MLC. Для подключения к серверу были использованы четыре из восьми портов 8 Гбит/с Fiber Channel.

Совместно с заказчиками был выбран профиль нагрузки — 70% чтения и 30% записи случайными блоками по 4К. Серверу было презентовано 15 томов, на каждый из которых подавалась выбранная нагрузка. В результате суммарное количество операций ввода-вывода составило более 250 тысяч, т.е. 1 Гбайт/с в терминах пропускной способности. Время отклика на чтение и запись оказалось одинаковым и составило менее 450 мкс. Далее мы, уменьшая нагрузки, устранили один из двух контроллеров, что не привело к падению времени отклика. Проверки удостоились и единственные движущиеся части системы — 2 из 6 вентиляторов также были устранены "на живую". Не привела к падению производительности и постоянная нагрузка в течение 5 дней.

Мы видим нишу СХД в ряде задач. Это масштабная виртуализация рабочих мест, в которой использование Violin Metro дает возможность разместить больше клиентов на 1 сервер, позволяя сэкономить на лицензиях гипервизора. Размещение ключевых приложений на подобную систему позволяет повысить производительность не на проценты, а в десятки раз за счет скоростной обработки всего массива данных. Серьезной проблемой на сегодня является обработка аналитических запросов в реальном времени. Мы располагаем результатами тестов, в которых нагрузка данных аналитики сокращается с 1 часа до 10 минут, а формирование OLTP отчета — с 24 до 3-х часов, что позволяет принять решение с учетом результатов аналитики в тот же рабочий день.

Результаты тестирования, представленные вендором, на ряде других нагрузок даны в табл. 2.

Табл. 2. Примеры эффективности использования массивов Violin на различных нагрузках.

| Область  | Тип нагрузки             | Влияние               | Результат  |
|--|--------------------------|-----------------------|--|
| Бизнес-аналитика (SAP)   | Комплексные запросы      | 4 мин → 3 сек         | 80x улучшение  |
| BI-нагрузка  | ETL                      | 1 час → 10 мин        | 6x улучшение   |
| VDI (VMware View, Intel Sandy Bridge)  | Рабочие места            | 10 → 25 польз./ядро   | 2.5x больше польз./ядро (2,560 ядер!)                          |
| OLTP-отчеты (SAP BW)   | Комплексные запросы      | 24 часов → 3 часа     | 8x улучшение   |
| CRM-система (Remedy)   | Пользовательские запросы | 8-10 мин → 30 сек     | 20x улучшение  |
| Биллинг (Keenan)   | Месячный биллинг         | 72 час → 22 час       | 330% улучшение   |
| Управление цепочками поставок (SAP Sales Distribution, Materials Management) | ERP                      | 7 час → 2 час         | 350% улучшение   |
| VDI (XenServer, Intel)   | Рабочие места            | 6 → 15 польз. На ядро | Меньше портов SAN (112) и серверов (56) для 3000 пользователей |

## EMC: флэш-продукты семейства Xtrem

Март 2013 г. — Корпорация EMC (NYSE:EMC) объявила о выпуске семейства оптимизированных для флэш-технологий серверных продуктов и СХД EMC® Xtrem™, а также анонсировала новую линейку флэш-карт EMC XtremSF™ на базе PCIe, которые существенно повышают производительность приложений.

XtremSF представляет собой аппаратный флэш-компонент для серверов, предлагаемый в виде широкой гаммы карт eMLC и SLC. Он может быть развернут как DAS-система, которая находится непосредственно на сервере и обеспечивает высокую производительность, или в сочетании с ПО для серверного кэширования EMC XtremSW™ Cache (прежнее наименование: EMC VFCache™) для радикального повышения производительности массива хранения с сохранением уровня защиты, требуемого для сред критически важных приложений. Флэш-карта XtremSF пополнила растущую линейку флэш-продуктов EMC для обеспечения сверхвысокой производительности, в которую входят флэш-оптимизированные гибридные массивы хранения EMC Isilon®, EMC VMAX® и EMC VNX®, а также горизонтально масштабируемый флэш-массив хранения EMC XtremIO (прежнее наименование: Project X).

Корпорация EMC также объявила о начале поставок XtremIO™ отдельным заказчикам. Массив XtremIO™ специально создан для использования преимуществ флэш-технологий и обеспечения реальной производительности, простоты администрирования и предоставления расширенных сервисов данных. Его горизонтально масштабируемая архитектура обеспечивает более высокий показатель «функциональных IOPS» для приложений, которым требуется высокий уровень производительности при произвольных операциях ввода-вывода, — например, для баз данных OLTP, виртуализации серверов и инфраструктуры виртуальных рабочих мест (VDI). Функциональные IOPS измеряются в условиях реальной эксплуатации, которые встречаются в современных требовательных производственных средах, когда все сервисы данных включены, а нагрузка инфраструктуры близка к предельной.

В таких реальных условиях система XtremIO демонстрирует показатели более 150 тыс. функциональных IOPS при смешанных операциях чтения/записи 4 КБ и 250 тыс. функциональных IOPS при операциях чтения 4 КБ для каждого блока X-Brick (блоки горизонтального масштабирования для массива XtremIO), а также показатели более 1,2 млн функциональных IOPS при смешанных операциях чтения/записи 4 КБ и 2 млн функциональных IOPS при операциях чтения 4 КБ в случае горизонтального масштабирования до кластера из восьми блоков X-Brick. Массив XtremIO обеспечивает такой уровень производительности со ста-

бильным временем отклика менее миллисекунды при работе с самым полным в отрасли набором интегрированных и флэш-оптимизированных сервисов данных, включая специфическую защиту данных для флэш-систем, "тонкое" выделение ресурсов, глобальное сокращение данных "на лету", ускоренное выделение ресурсов VMware через VAAI и снимки файловой системы с правом записи.

### Основные характеристики технологии XtremSF

Серверные флэш-карты PCIe семейства XtremSF представляют заказчикам следующие преимущества:

- **лидирующие в отрасли показатели производительности.** По результатам проведенных испытаний, флэш-устройства XtremSF способны обеспечивать рекордный показатель 1,13 млн IOPS в стандартном формфакторе;
- **беспрецедентную гибкость.** Флэш-устройства XtremSF предлагаются в широком диапазоне емкости на базе eMLC (550 ГБ и 2,2 ТБ) и SLC (350 ГБ и 700 ГБ). Кроме того, в случае развертывания совместно с интеллектуальным ПО для серверного флэш-кэширования XtremSW Cache флэш-карты XtremSF могут использоваться как устройства кэширования для повышения производительности с защитой массива для таких приложений, как Oracle, Microsoft SQL и Microsoft Exchange;
- **новый уровень эффективности.** Флэш-карты XtremSF обеспечивают снижение совокупной стоимости владения на 58% по сравнению с другими продуктами. Все продукты XtremSF (включая исполнение емкости 2,2 ТБ) представляют собой карты PCIe 25 W половинной длины и половинной высоты, что позволяет получить максимальную емкость хранения в минимальном пространстве, наилучшие показатели плотности и минимальные затраты на энергопотребление, снизив при этом коэффициент использования ЦП практически вдвое (снижение до 50%).

### Стратегия EMC Flash Software

В 2012 г. корпорация EMC выпустила флэш-карту XtremSW Cache (прежнее наименование: VFCache). Этот первый шаг в реализации долговременной стратегии EMC по разработке серверных флэш-технологий позволил получить серверную СХД, в основу которой положено ПО, работающее на флэш-устройствах PCIe на базе SLC. Это ПО превращает серверную флэш-систему DAS в устройство кэш-памяти и позволяет повысить производительность самых различных критически важных приложений, обеспечивающих выполнение транзакций и необходимых для принятия решений. Теперь ПО XtremSW Cache может использоваться с различными флэш-картами XtremSF eMLC и SLC типов.

В будущем EMC планирует выпустить EMC XtremSW Suite — комплексный пакет ПО для флэш-систем, не зависящий от конкретных устройств. Этот новый программный пакет не только будет пре-