

RAID-контроллеры с поддержкой функции HBA

Последовательное развитие семейств продуктов Adaptec — адаптеров HBA и RAID-контроллеров, последние анонсы которых прошли весной и в начале сентября 2013 г., позволило реализовать поддержку интерфейса 12 Гбит/с, а также значительно расширить функционал этих решений (maxCache Plus, Adaptec RAID Code, Flexible Configuration, Hybrid RAID и Optimized Disk Utilization).



Дмитрий Зотов — инженер компании PMC-Sierra, подразделение Adaptec by PMC.

HARDWARE vs. SOFTWARE RAID. Окончание великой эпохи противостояния

Современные RAID-контроллеры незаметно стали применять несколько другой подход к обоим решениям: software и hardware RAID, отличный от того подхода, который доминировал еще буквально 1–2 года назад.

В силу этого спорить о том, что лучше: HARDWARE или SOFTWARE RAID-решение — в большинстве случаев не имеет большого смысла, так как можно легко испробовать оба решения на практике, причем одновременно на одном и том же контроллере, не меняя ничего в системе, кроме настроек, и придти к выводу, что же вас больше устраивает. Исключения составляют HBA-платы, не поддерживающие режим RAID (рис. 1).

Рассмотрим возможности использования режимов БЕЗ функции RAID (режим HBA, как некий начальный уровень для software RAID) в текущей линейке HBA и RAID-контроллеров Adaptec (табл. 1).



Рис. 1. SAS HBA Adaptec 71605Ne с поддержкой до 128 дисков и возможностью шифрования (нет функции RAID).

Табл. 1. Поддержка режимов RAID/HBA в семействе контроллеров Adaptec.

Модели из текущей линейки HBA/RAID-контроллеров Adaptec	Поддержка режима HBA	Поддержка режима RAID
6H/7H/7Ne HBA	Да	—
RAID-контроллеры Series 6 (включая 6E и 6Q)	Нет	Да
RAID-контроллеры Series 7 (включая 7E и 7Q)	Да	Да
RAID-контроллеры Series 8 (включая 8Q)	Да	Да

Непривычно, но факт — RAID-контроллер и HBA “в одном флаконе”. Итак, на одном и том же контроллере можно легко использовать часть дисков в режиме HBA (при этом часть из них можно отдать для использования в software RAID), а часть дисков — в составе RAID-тома, созданного средствами самого RAID-контроллера (рис. 2).

Как видно на рис. 2, часть дисков в режиме HBA передается операционной системе и работа с ними выглядит так же, как и с дисками, подключенными к портам HBA, интегрированным на материнскую плату. Еще часть дисков передается ОС для создания RAID-тома F: средствами операционной системы (software RAID). Важно отметить, что диски могут переда-

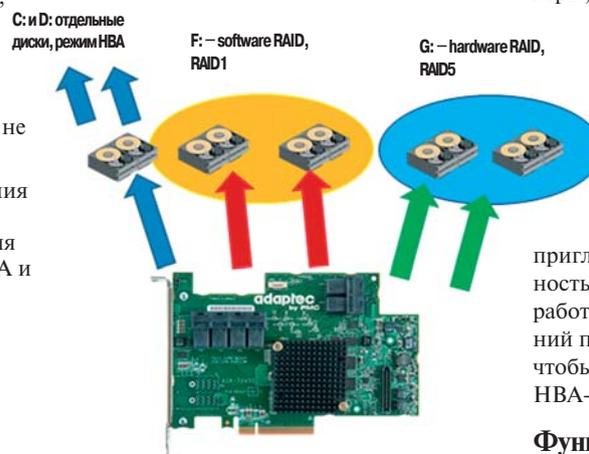


Рис. 2. Возможность использования режима HBA, решений типа SOFTWARE RAID и HARDWARE RAID на одном контроллере (software RAID создается и поддерживается средствами операционной системы).

ваться не только в режиме HBA, но и в других режимах, например, Simple Volume. Оставшиеся диски с помощью RAID-контроллера образуют том G: (hardware RAID).

Использование режима HBA

Режим HBA, при использовании на RAID-контроллере, позволяет расширить границы практического применения RAID-контроллера. Примеры практического применения режима HBA для продуктов типа HBA-карточка и RAID-контроллер с поддержкой режима HBA даны в табл. 2.

Уголок профессионала: режимы создания Software RAID

Варианты реализации режима software RAID с преимуществами и недостатками приведены в табл. 3.

Новые функции RAID-контроллеров с поддержкой режима HBA

Появление контроллеров, одновременно поддерживающих режим HBA и RAID, привело к изменению ряда подходов и настроек и появлению новых функций.

Кроме того, с учетом эволюции и развития таких решений, как RAID-контроллеры, также появляются новые функции в стеке, не связанные напрямую с режимом HBA, например, поддержка дисков 4K native и поддержка uEFI. Рассмотрим некоторые их особенности.

В настройках контроллера (это характерно для Series 7 и Series 8 контроллеров) под MBR BIOS, по приглашению Ctrl+А, исчезла возможность назначать HOT SPARE диски. Разработчики пошли на это в силу ограничений памяти на начальном этапе загрузки, чтобы добавить функции обслуживания HBA-режима.

Функция Uninitialize

В контроллерах Series 7 и Series 8 появляется необычная функция Uninitialize. Мы получаем два пункта меню Initialize и Uninitialize.

Табл. 2. Режим HBA, при использовании на RAID-контроллере, позволяет расширить границы практического применения RAID-контроллера.

Режим	Поддержка	Комментарии
Передача полной емкости диска в OS в режиме HBA	HBA и RAID-контроллеры с режимом HBA 6H, 7H, 7He, Series 7 и Series 8 RAID-контроллеров Adaptec	Как правило, не является решением для хранения пользовательских данных серверной системы в силу низкой надежности. Используется для создания служебных томов
Передача полной емкости диска в OS в режиме HBA для создания SOFTWARE RAID	HBA и RAID-контроллеры с режимом HBA 6H, 7H, 7He, Series 7 и Series 8 RAID-контроллеров Adaptec	Режим активно используется для создания устройств хранения или пользовательских томов высокой производительности
Поддержка недисковых устройств в режиме HBA	HBA и RAID-контроллеры с режимом HBA 6H, 7H, 7He, Series 7 и Series 8 RAID-контроллеров Adaptec	Удобная возможность поддержки нужного количества недисковых SAS/SATA устройств, например, устройства типа SATA CD-ROM
Поддержка сложных недисковых устройств с помощью функции MultiLUN	HBA и RAID-контроллеры с режимом HBA 6H, 7H, 7He, Series 7 и Series 8 RAID-контроллеров Adaptec	Поддержка сложных недисковых устройств, типа ленточных библиотек
Полная поддержка внешних RAID-стоек (RBOD) ¹⁾	Series 8 RAID-контроллеров Adaptec	Поддержка внешних RAID-стоек с возможностью создания какого-то количества RAID-томов на этих стойках и передачи этих томов в операционную систему.

¹⁾ поддержка данных функций может измениться в самое ближайшее время с выходом новых прошивок для HBA и RAID-контроллеров

Табл. 3. Варианты реализации режима software RAID.

Решение	Преимущества	Недостатки
Software RAID поверх Hardware RAID	Не имеет особо значимых преимуществ в области производительности и надежности	Структура крайне сложная, в силу этого НЕ ИМЕЕТ ПРАКТИЧЕСКОГО ПРИМЕНЕНИЯ
Software RAID поверх Simple Volumes	Возможность использования кэш контроллера при работе с Simple Volume. Возможность создавать Simple Volume тома нужной емкости (меньше емкости отдельного диска)	Совместимость только внутри данного стека от данного производителя. Например, если такое решение создано на Adaptec RAID 5405, то жесткие диски такого тома можно перенести только на семейство контроллеров с таким же RAID-стеком (на контроллеры Series 2,3,5,6,7). Невозможно использовать HBA, как отдельные карточки или интегрированные на материнские платы HBA для поддержки этого решения
Software RAID поверх режима HBA	Высокая совместимость решения. Диски (это зависит от типа дисков SAS или SATA, SATA наиболее универсальны) могут переноситься на любой SAS/SATA HBA, интегрированный на материнская плату и отдельный, любого производителя	Можно выделить много недостатков такого решения, которые можно устранить тем или иным способом. В качестве примера можно указать сложности с организацией и поддержкой ЗИП пула жестких дисков для ЦОД (см. главу, посвященную Coercion Factor)

Если функция Initialize создает специальную структуру данных на диске, называемую METADATA, то функция Uninitialize убирает эту структуру с диска.

Область метаданных – служебная область, которая недоступна для операционной системы. В области метаданных хранится важная информация о структуре физических дисков, партиций, RAID-томов и т.д.

Совместно с функцией Initialize в атоматическом режиме выполняется и другая функция – Coercion. Назначение – небольшая “подрезка емкости” для удобной замены дисков одного производителя дисками другого.

Табл. 4. Алгоритм работы функции Coercion для дисков архитектуры 512K.

Drive Size	Coercion Factor
< 256GB	100MB
>= 256GB and < 1TB	1GB
>= 1TB	5GB

Алгоритм работы функции Coercion для дисков архитектуры 512K представлен в табл. 4.

Coercion выполняется автоматически при инициализации жесткого диска, применении режима JBOD (метаданные резервируются, но не используются) и при назначении Hot Spare.

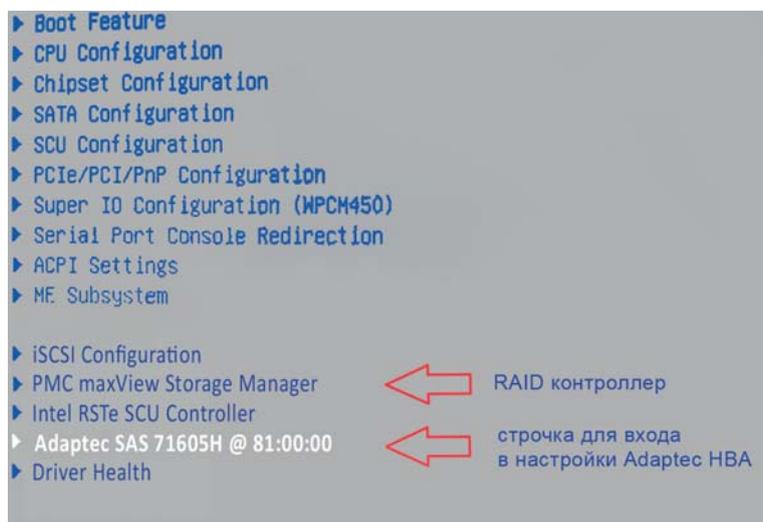


Рис. 3. Особенности настройки RAID-контроллеров и HBA в режиме uefi.

adaptec
by PMC

RAID контроллеры Series 8



Производительность 12 Гб/с
Плотность компоновки
Уровневое управление

- Семейство продуктов 12 Гб/с для PCIe 3.0
- Производительность > 700 тыс. IOPS
- 16 и 8 внутренних и внешних портов в компактном форм-факторе LP/MD2
- Adaptec maxCache Plus для кэширования и уровневого управления
- Интегрированная защита кэша на базе флэш-памяти

maxCache Plus



Для приложений, требующих высокой производительности и плотности компоновки

Важно. При выполнении Uninitialize автоматически выполняется “UNCOERCION” (освобождение емкости) с возможностью ее использования для размещения пользовательских данных.

После Uninitialize диски доступны для менеджера управления системой хранения (disk manager) операционных систем в режиме HBA. Инициализированные диски, при использовании функции initialize, не доступны OS пока из них не будет создан объект уровня RAID-том (JBOD, SIMPLE VOLUME, RAID1 и т.д.).

Важно. HBA-продукты и HBA-режимы RAID-контроллеров не используют Coercion.

Функции загрузки

Режим HBA подразумевает другой механизм загрузки. WWN-адрес загрузочного диска хранится в памяти контроллера, в то время как для RAID-режима информация о загрузке, буквально хранится в области Metadata (метаданных).

Поддержка uEFI

Важно. Если ваша серверная плата поддерживает режим uEFI, настройки RAID- и HBA-контроллеров производятся через системный BIOS (BIOS материнской платы), традиционное приглашение в настройках BIOS-контроллера недоступно. Обычно они помещаются в меню “Advanced” (рис. 3):

- для RAID-контроллеров — строка “maxView Storage Manager”;
- для HBA — строка выглядит, как “Adaptec SAS HBA-модель”.

Подробности настройки и обслуживания функций загрузки для режимов HBA и RAID можно найти в User Guide и Read_Me документах на сайте поддержки www.adaptec.com/support.

Более детальная информация о функциях на RAID- и HBA-контроллерах доступна на сайте ask.adaptec.com.

*Дмитрий Зотов,
компания PMC-Sierra,
подразделение Adaptec by PMC*

EMC: начало продаж XtremIO™

Ноябрь 2013 г. — Корпорация EMC объявила о начале продаж EMC® XtremIO™ — массива на флэш-дисках, который обеспечивает стабильно высокую прогнозируемую производительность для любых рабочих нагрузок за любой период времени независимо от статуса и заполнения массива. Этого удалось достичь благодаря использованию в EMC XtremIO нескольких уникальных инноваций в области флэш-технологий: горизонтально масштабируемой многоконтроллерной архитектуре с линейной масштабируемостью, встроенной и постоянно действующей дедупликации и защите данных, превосходящей в 6 раз по эффективности и в 4 раза по производительности традиционные RAID.

По прогнозам IDC, объем рынка флэш-массивов к 2015 г. составит \$1,2 млрд (“Worldwide Enterprise All-Solid State Storage Array 2013—2016 Forecast”, IDC, апрель 2013 г.).

Основные технологические характеристики XtremIO

Заказчикам необходимы флэш-массивы для поддержки рабочих нагрузок в тех случаях, когда требуется обеспечить прогнозируемое и стабильное малое время ожидания при работе с часто меняющимися наборами данных — инфраструктурами виртуальных рабочих мест (VDI), виртуальными серверами, крупными консолидированными базами данных и средами тестирования и разработки. Благодаря XtremIO обеспечивается не только лучшая производительность для этих рабочих нагрузок, но и снижение стоимости IOPS, а также простота администрирования. Четыре особенности архитектуры принципиально отличают XtremIO от других массивов на флэш-дисках:

- *размещение данных на основе содержания* поддерживает внутреннюю балансировку и оптимизацию массива до долей процента для всех твердотельных дисков (SSD) и контроллеров массива, а также удаляет дублированные данные “на лету” на пути передачи данных;
- *двухэтапный механизм обработки метаданных* позволяет массиву полностью использовать присущие флэш-технологии преимущества случайного доступа и размещать данные в любом месте массива без применения внутренних процессов чистки (называемых также “сбор мусора”). Это позволяет XtremIO избежать свойственных другим флэш-массивам снижения показателей производительности в IOPS на 50%, повышения задержки на 1000% и 10-кратного сокращения срока службы флэш-дисков;
- *алгоритм защиты данных XDP (XtremIO Data Protection)* — это специфичный для флэш-технологий алгоритм, который обеспечивает защиту от сбоя твердотельных дисков (SSD) и позволяет в 6 раз повысить полезную емкость по сравнению с классическими RAID. Алгоритм XDP позволяет конечным пользователям использовать 100% емкости XtremIO, сохраняя максимальный уровень производительности. В других флэш-массивах снижение производительности начинается при заполнении массива на 60—80% емкости. Таким образом, использование алгоритма XDP на 40% повышает полезную емкость массива по сравнению с другими флэш-массивами. Кроме того, при длительной эксплуатации в условиях ЦОД эффективные алгоритмы XDP обеспечивают четырехкратное повышение производительности и срока службы флэш-дисков по сравнению с RAID;
- *совместно используемые метаданные (shared in-memory metadata)* позволяют массиву обеспечивать широкий спектр показателей производительности и быстро клонировать хранимую в массиве информацию, что обеспечивает резкое ускорение выполнения таких стан-

дартных задач, как развертывание виртуальных машин и т.п. Клонирование виртуальных машин производится со скоростью до 20 раз превышающей ширину полосы пропускания между хостом и массивом, выполняется в несколько раз быстрее и с меньшим влиянием на производственные виртуальные машины, чем в других массивах на флэш-дисках.

XtremIO — это горизонтально масштабируемый массив, построенный из строительных блоков, называемых «X-Brick». Для каждого модуля X-Brick предусмотрена емкость 10 ТБ; дополнительный вариант с емкостью 20 ТБ планируется выпустить в начале следующего года. Массивы XtremIO демонстрируют производительность в 1 млн случайных операций ввода-вывода в секунду (IOPS) с эффективной емкостью 250 ТБ в одном кластере XtremIO с встроенной дедупликацией. Один кластер XtremIO масштабируется с двух до восьми контроллеров и до 128 ядер и может работать с любой базой данных OLTP, виртуальным сервером и инфраструктурой VDI со всеми активными сервисами данных.

Управление массивом XtremIO также интегрировано с VMware vSphere и ускоряется при помощи программных интерфейсов (API) VMware VAAI. В дополнение к этому, XtremIO совместим с другими технологиями EMC, включая EMC VPLEX™, EMC PowerPath® и службу EMC Secure Remote Support (ESRS).

Мнения заказчиков и партнеров

*Расс Норман (Russ Norman), старший менеджер по системному инжинирингу компании E*TRADE*

«Производительность, надежность и скорость — вот критически важные факторы для нашей среды приложений. Нам требуется инфраструктура, которая обеспечивает очень низкое время задержки для быстрой обработки заказов. Мы развертываем массив XtremIO не только для того, чтобы выйти на новый уровень производительности, но — что еще важнее — чтобы перейти на новый уровень использования наших приложений. Благодаря низкому времени задержки хранения мы можем обрабатывать больше операций ввода-вывода в секунду и перемещать большие объемы данных. Мы определили, что в среднем время отклика для записи в массив XtremIO составляет от 0,25 до 0,3 мс, а отклик при чтении колеблется где-то вокруг отметки 0,2 мс. Причем это стабильные показатели при рабочих нагрузках между 100 тыс. и 200 тыс. IOPS. Кроме того, для нас важными аргументами в пользу массива XtremIO с точки зрения эффективности затрат и емкости стали “тонкое” выделение ресурсов и встроенная дедупликация”.

Тодд Павонн (Todd Pavone), исполнительный вице-президент по разработке и стратегии продуктов компании VCE.

“Специализированная система Vblock для экстремальных нагрузок с флэш-массивами EMC XtremIO обеспечивает исключительно высокую производительность для масштабируемых решений на базе инфраструктуры виртуальных рабочих мест (VDI) для работы с большими объе-