

HP: НОВЫЙ ФЛЭШ-МАССИВ

В июне 2013 г. компания HP сделала ряд важных анонсов, среди которых: анонсирование доступности HP ZPAR StoreServ 7450 – оптимизированного массива, составленного только из флэш-накопителей (SSD-накопителей) с общей производительностью более 500 000 операций ввода-вывода в секунду и задержками менее 0,7 мс на операцию ввода-вывода.



Андрей Вересов – менеджер по развитию бизнеса систем хранения данных, HP в России.

Введение

В июне 2013 г. компания HP анонсировала модели систем хранения среднего класса, оптимизированные на работу с флэш-носителями – ZPAR StoreServ 7450 (рис. 1). И теперь практически любая система хранения HP, представленная на рынке, поддерживает накопители SSD.

Традиционно, требование высокой производительности от дисковой подсистемы достигалось добавлением большого количества вращающихся шпинделей, как правило, с минимальной утилизацией их объема (рис. 2). Использование SSD-накопителей позволило сократить число вращающихся шпинделей, подняв при этом производительность массивов данных.

HP ZPAR 7450 с SSD

В то время когда на рынке доминируют традиционные массивы, имеющие почти 20-летнюю архитектуру, либо сверхдорогие all-flash стартапы, HP предлагает массив с совершенно другой архитектурой, специально разработанный для одновременной работы с различными типами данных. Ключевые особенности нового массива ZPAR 7450 заключаются в том, что контроллеры не являются узким местом в производительности, в связи с чем достигается практически ее линейный рост по мере добавления новых дисков.

Технологически эти массивы схожи на ZPAR 7400 с улучшенными характеристиками кэша и интерконнект-связей между контроллерами. Как и модель 7400, новый массив 7450 может расширяться до массива из 4 полностью связанных контроллеров с зеркалированием кэша на запись. При этом уникальность массива состоит в том, что каждый его контроллер является активным для всех томов, что дает честную балансировку нагрузки между всеми нодами и всеми дисками массива. Подобная технология позволяет сохранять высокую производительность даже при выходе из строя одного контроллера.

Массивы 7450 имеют большой объем кэша – 48 GB для двухконтроллерной системы и 96 GB – для четырехконтроллерной, при этом кэш на запись имеет свои собственные каналы (32 GB/64 GB для каждой системы соответственно). Таким образом, операционная система массива не может "отьедать" кэш на запись под свои нужды, и он полностью доступен для операций ввода/вывода.

Благодаря такому большому уровню кэша, использованию чипов ASIC и восьмиядерных процессоров Intel E5-2470 2.3GHz массив может разогнаться до 520 000 IOPs при использовании только SSD-дисков (4k random read), что важно при задержке менее 1 мс. Большие IOPS требуют большой пропускной способности массива. Благодаря использованию чипов ASIC значительно разгружается центральный процессор каждого контроллера. Это позволяет получить очень высокие показатели пропускной способности – до 5.2 GB/s. Таким образом, "сырая" вычислительная мощность каждого контроллера возросла более чем на 35%.

Чтобы справляться с такими большими нагрузками и поддерживать кластеризацию, была увеличена пропускная способность шины между контроллерами.

Поскольку большинство традиционных дисковых массивов построены на архи-



Рис. 1. Позиционирование массивов HP ZPAR StoreServ 7450 в семействе СХД HP ZPAR.

тектуре 20-летней давности, то использование SSD-дисков в них не приносит существенной производительности, а лишь приводит к кратковременному улучшению показателей чтения. Контроллеры таких массивов имеют существенные ограничения по количеству поддерживаемых SSD-накопителей.

Особенности архитектуры массива ZPAR таковы, что позволяют получить высокую производительность при добавлении большого количества SSD (120 SSD – для модели ZPAR 7200, 240 SSD – для моделей ZPAR 7400/7450).

Так же, как и модели 7400, модель ZPAR flash-optimized 7450 поддерживает до 480 дисков LFF/SFF. Для оценки износа SSD-дисков используется продвинутый механизм Flash Wear Gauge, который во время проинформирует о возможных проблемах с носителями.

Для массива ZPAR 7450 был переработан микрокод по работе с кэшем контроллера.

Минимальный объем страницы, хранимой в кэше большинства систем хранения, – 16 Кбайт. У массива ZPAR "умный" кэш – каждому типу дисков, в зависимости от их объема и количества, выделяется определенная область кэша на запись. Дальше, если к этому типу дисков параллельно будут идти последовательные и случайные обращения, то страницы этой части кэша адаптируются под типы запросов. Например, при нескольких последовательных операциях на чтение кэш сразу считывает несколько блоков по 16 Кбайт, таким образом снижается нагрузка на носители и уменьшается время отклика. Как только последовательные обращения прекращаются, в кэше остаются только актуальные страницы (рис. 3).

Рассмотрим пример, когда хост запрашивает 4KB на чтение. Для традиционных массивов от кэша к диску характерны обращения блоками в 16KB, поскольку у вращающихся дисков существует задержка. В связи с этим, в кэш помещается

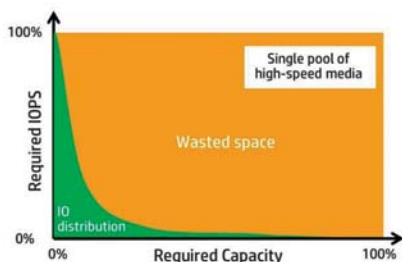
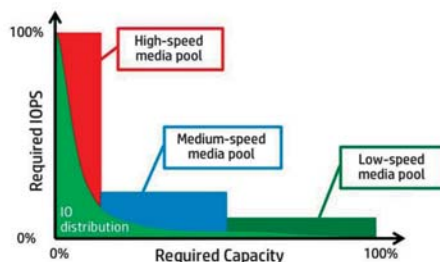


Рис. 2. Достижение высокой производительности в массивах традиционного типа и в современных массивах.



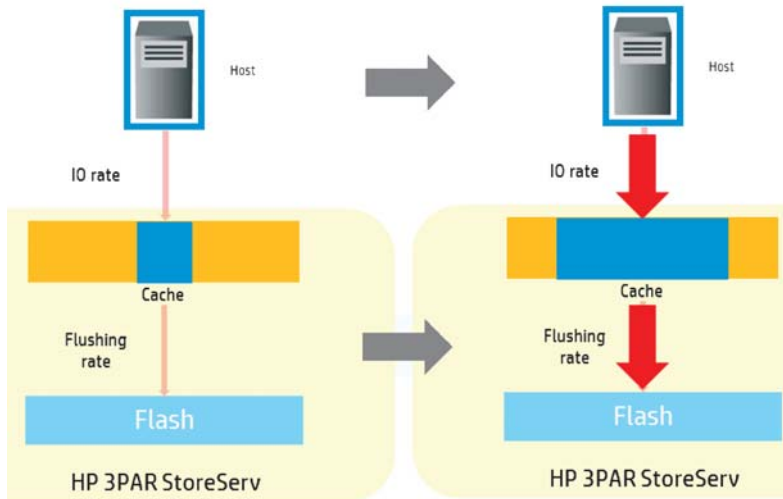


Рис. 3. Адаптивный кэш в системе хранения 3PAR 7450.

сразу блок большего размера, чтобы прогнозировать дальнейшие обращения к соседним блокам. Такой подход повышает скорость обращения к массиву, но сильно загрузит кэш данных. Для SSD-дисков характерным является время отклика, меньшее в десять раз. Поэтому в массивах 3PAR 7450 применяется адаптивный способ работы: кэш массива запрашивает от SSD-носителей блок, практически равный по размеру блоку I/O, запрошенному от хоста, что меньше нагружает кэш и SSD-диски (рис. 4).

Запросы на запись усложняют ситуацию. Известно, что SSD-носители обрабатывают запись блоками размером 4/8 KB от хоста, но при этом перезапись ячеек SSD-носителя идет блоками по 64KB, для ранних моделей накопителей SSD – 512KB (циклы перезаписи для SSD всегда лимитированы). Подход традиционных массивов: в кэше хранятся страницы по 16KB и на SSD-носители записываются блоки такого же размера, даже если сервер отправляет блоки по 4/8KB, т.е. неполная страница будет храниться в кэше до заполнения, только после этого она будет записана на SSD-носители. Но чаще всего на SSD-носители будет сбрасываться неполная страница, т.к. в массивах применяются политики принудительного сброса кэша на диски по достижении определенного тайм-аута для сохранения данных

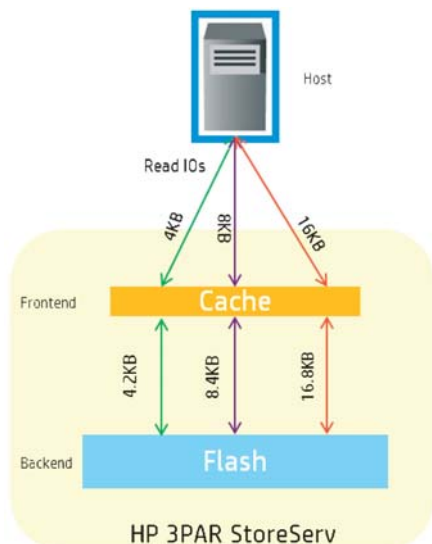


Рис. 4. Принцип работы кэша на чтение в массиве 3PAR 7450.

(даже если в кэше контроллера записано только 4KB, то сбрасывается блок в 16KB, из которых 12KB будут заполнены нулями). В таком подходе эффективность использования SSD-носителей и срок их службы сильно снижаются. Такое поведение характерно для большинства массивов традиционного типа.

Подход, используемый в системах хранения 3PAR, – на диски с кэша сбрасываются блоки меньшего размера (до 4KB) – имеет ряд преимуществ:

- уменьшается износ SSD-носителей, т.к. записываться будут только блоки с данными, нулевые блоки записываться не будут;
- улучшаются механизмы поиска свободного места на SSD после заполнения носителя: массиву легче найти свободный блок в 4KB, чем блок в 16KB;
- с учетом предыдущего преимущества улучшается утилизация SSD-носителя – хранятся только блоки с данными; блоки, заполненные нулями исключены;
- для томов RAID1 уменьшаются задержки и штрафы, связанные с последовательностью запросов на "чтение – изменение – запись" (рис. 5).

Если же от хоста идут запросы на запись большими блоками (например, 128KB), то кэш массива 3PAR разбивает их на блоки меньшего размера и разделяет нагрузку между носителями. Помимо этого, контроллер умеет одновременно обрабатывать запросы разных типов: последовательные и случайные. Каждый тип запроса параллельно обрабатывается чипами ASIC, тем самым снижается очередь и повышается скорость работы массива в смешанных нагрузках. В массивах традиционного типа на смешанных нагрузках характерны большие очереди запросов (рис. 6).

Какие преимущества дает СХД 3PAR 7450 в сравнении с проектами flash-массивов, появившимися в последнее время?

Во-первых, массив 3PAR 7450 дает высокую производительность, сопоставимую с flash-массивами. При этом используются SSD-диски, кото-

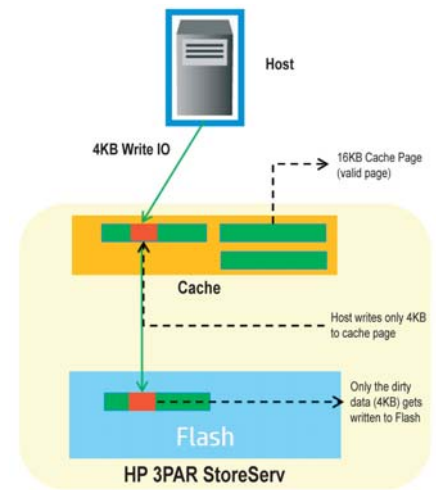


Рис. 5. Принцип работы кэша 3PAR 7450 на запись.

рые значительно дешевле flash-модулей. Традиционные системы хранения данных, в силу устаревшей архитектуры, не могут добиться высоких показателей производительности.

Во-вторых, массив 3PAR поддерживает Disaster recovery, что пока не могут предложить производители flash-массивов.

В-третьих, flash-массивы, как и массивы традиционной архитектуры, не могут наращиваться до 4-х контроллеров с автоматической балансировкой нагрузки между ними.

В-четвертых, сам по себе flash-массив обычно слабо связан с другими системами хранения, что уменьшает зону для его применения, в то время как 3PAR позволяет хранить все типы данных на одной системе хранения, с одним интерфейсом управления и динамическим перемещением данных между уровнями.

В-пятых, массивы 3PAR могут объединяться в федерацию из нескольких систем, с балансировкой нагрузки между массивами, что пока недоступно как в решениях flash, так и в массивах традиционной архитектуры.

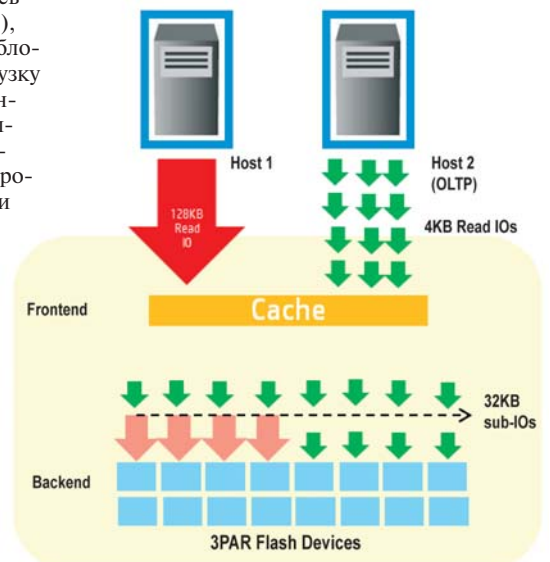


Рис. 6. Принцип работы кэша массивов 3PAR в смешанных нагрузках.

Поддержка Disaster Recovery

Помимо улучшений в области производительности, в массивах ZPAR возросли возможности по созданию отказоустойчивых решений. В частности, усовершенствовалась технология Peer Persistence, направленная на совместную работу с решениями VMware по следующим направлениям:

- реализованы функции балансировки нагрузки между двумя площадками совместно с Storage vMotion. Теперь автоматическая балансировка нагрузки VMware возможна не только внутри массива между всеми контроллерами, но и между несколькими массивами;
- реализована функция High Availability для виртуальных машин между несколькими датацентрами;
- синхронная репликация теперь поддерживает технологию "непропадающего тома" (never disappearing LUN presentation) — том реплицируется между массивами и для него создаются отказоустойчивые пути, презентованные хосту. Дисковые массивы, размещенные на разных площадках, будут связаны между собой репликацией, реализованной при помощи ZPAR Remote Copy. Ввиду того, что обе площадки являются активными, репликация будет двусторонней. Технология ZPAR Peer Persistence предоставляет узлам кластера vMCS равномерный доступ к двум системам хранения на площадках датацентра, при этом каждый логический диск (LUN) каждой системы хранения ZPAR имеет синхронную копию на другой площадке (рис. 7). Данная копия представляется виртуальной среде VMware пассивным путем к основному LUN, так как имеет такой же идентификатор WWN (уникальный идентификатор устройства в сетях SAN). В случае сбоя основного LUN, его дублер остается видимым для хостов, таким образом обеспечивая возможность гипервизора перенести виртуальные машины с использованием обычных технологий vMotion, High-Availability или Fault-Tolerance.

Большинство традиционных технологий (таких, как например, автоматизированное решение VMware Site Recovery Manager) в случае сбоя системы хранения данных может перевести всю виртуальную среду на резервную площадку, но для этого будет необходима перезагрузка виртуальных машин. Благодаря же вышеописанной технологии Peer Persistence, для виртуального кластера система хранения доступна даже тогда, когда дисковый массив одной из площадок вышел из строя. В то же время, переключение на резервную площадку происходит без использования сторонних механизмов, что дает возможность перевести виртуальные машины на другую площадку без их останова и, при этом, не усложняет управление инфраструктурой (см. рис. 7).

Для физических машин по-прежнему доступно ПО кластеризации приложений, работающее совместно с массивами ZPAR — HP Cluster Extension для Linux и Windows (<http://h18006.www1.hp.com/storage/software/ces/index.html>) и MetroCluster для HP-UX.

Система шифрования данных в массивах ZPARStoreServ 7000/10000

С июня этого года у заказчиков появилась возможность заказывать, помимо стандартных массивов ZPAR 7000/10000, массивы с функцией шифрования данных. Отличие этих массивов — в наличии специальных дисков SED (Self-Encrypted Disk), в каждом из которых есть чип шифрования. Поддерживаются ключи шифрования AES 256 bit и 32 byterandom. Благодаря низкой задержке чипа, скорость работы SED-диска практически не отличается от скорости обычного диска. При первой инициализации ключ шифрования распознается менеджером LocalKeyManager, который является частью операционной системы массива и выполняется на каждом контроллере. Во время первой инициализации массивом будет предложено сохранить копию ключа. В случае пропадания питания на дисках или массиве данные блокируются, и доступ к данным может быть возобновлен при повторном вводе ключа шифро-

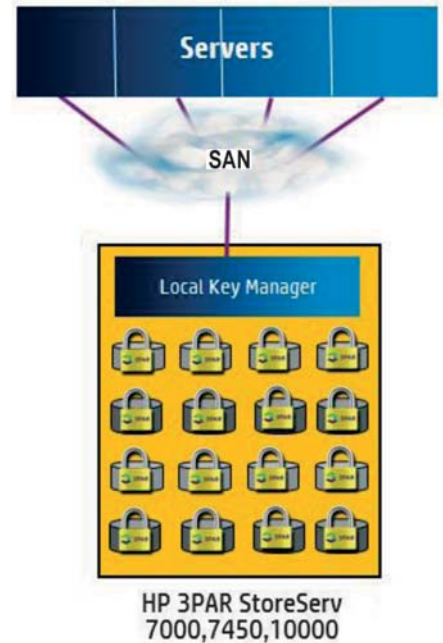


Рис. 8. Схема работы шифрования данных в массиве ZPAR.

вания в консоли Local Key Manager. Если IT-специалист уходит из организации, то можно поменять ключ. При миграции на массив без функции шифрования данные будут расшифрованы, при миграции в обратную сторону — зашифрованы. Есть важные нюансы:

- шифрование активируется лицензией, приобретается одна лицензия на массив;
- смешивать обычные и SED-диски в массиве с активированной функцией шифрования нельзя;
- если функция шифрования не активирована, то смешивать диски в массиве можно, но стоимость SED-дисков выше примерно на 5%;
- после того как функция шифрования активирована — все SED-диски должны быть переинициализированы, т.е. данные на дисках будут недоступны.

Для шифрования будут доступны несколько типов дисков SED — 400GB MLC SSD, 450GB 10k, 900GB 10k, 1TB 7.2k 2.5" (рис. 8).

Экосистема хранения данных нового поколения

Совсем недавно HP анонсировала поддержку Smart-кэша на хостах HP ProLiant (http://h18004.www1.hp.com/products/quickspecs/14532_div/14532_div.PDF). Суть этой технологии — в выделении группы дисков на сервере для хранения наиболее востребованных данных. В данной версии поддерживаются серверы HP ProLiant Gen8 с контроллерами HP Smart Array P22x, P42x, P721m, P822, P83x (необходим flash-модуль минимум 1GB). Пока поддерживаются 32 тома для кэширования данных. Следующим этапом станет расширение функционала этой технологии — на стороне хоста будет устанавливаться драйвер, поддерживающий тома системы хранения ZPAR для кэширования. Это позволит перемещать наиболее часто востребованные данные для чтения в кэш, созданный из группы дис-

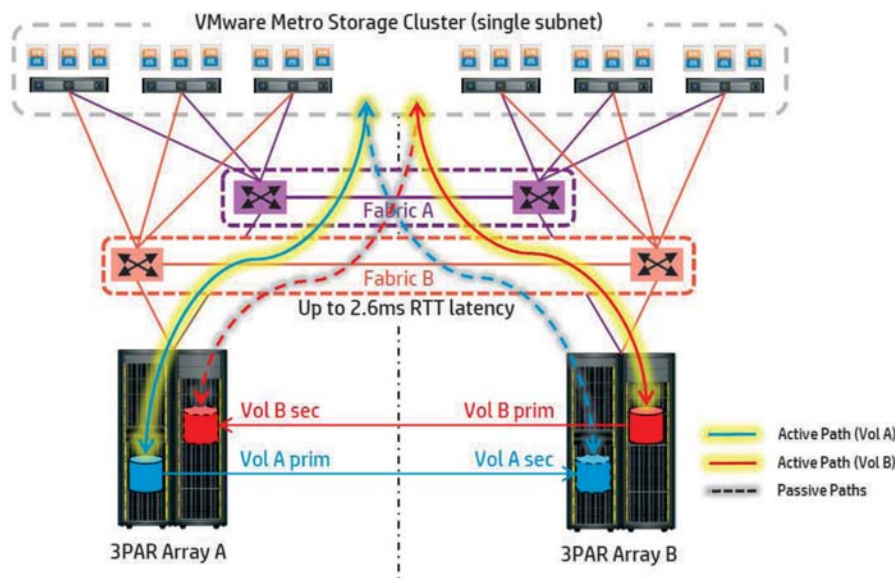


Рис. 7. Технология Disaster Recovery в среде VMware.

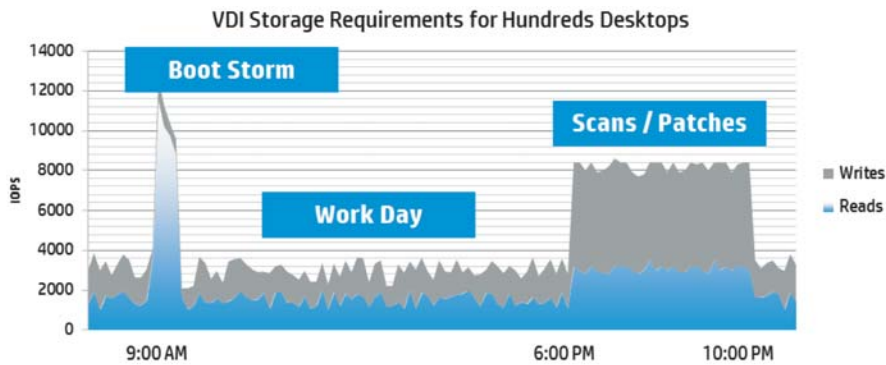


Рис. 10. Паттерн типичных обращений к массиву для VDI.

ков сервера. Кроме того, будут поддерживаться как SSD-носители, так и IO Accelerator'ы на шине PCI-E, установленные в сервер. Таким образом, для критичных приложений это позволит существенно поднять производительность на чтение (до 3-х раз), сократив задержки на доступ к данным.

Smart Cache поддерживает кластеризацию, интеграцию с vCenter посредством использования агентов для виртуальных машин (рис. 9).

Тем не менее, HP не предлагает перемещать все данные на flash, для некоторых приложений это неприемлемо, а большинству приложений быстрый SSD-уровень требуется только для кратковременного хранения данных, поэтому "закат эры HDD" пока не предвидится.

Для каких сред предназначен массив HP 3PAR 7450?

Во-первых, это технологии OLTP / WEB 2.0, для которых время отклика и скорость массива имеет решающее значение, а для предприятия это напрямую влияет на полученную прибыль.

Во-вторых, это ставшие широко востребованными технологии VDI, для которых характерно наличие лавинообразного роста обращений к массиву в течение очень короткого промежутка времени (bootstorm). Использование быстрого массива 3PAR 7450 позволит выжать максимум производительности из SSD-накопителей (до 10000 IOPS с носителя) и таким образом сократить их количество (рис. 10).

В-третьих, это бизнес-аналитика. Быстрый анализ большого количества данных — быстрая реакция на изменения и высокая прибыль.

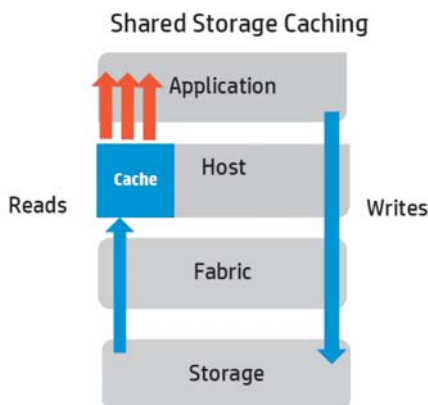


Рис. 9. Принцип работы SmartCache.

Заключение

В заключение стоит упомянуть также о следующих расширениях:

- *теперь можно заказывать MLC SSD-диски 400GB — для тех случаев, когда необходимы SSD-тома больших размеров, при этом необходимо помнить, что число циклов перезаписи у MLC-дисков существенно меньше, чем у SLC, NLSAS-диски 1TB 7.2k 2.5" — для тех, кому необходимы емкие хранилища в форм-факторе 2.5" (для 3PAR 7000), SAS-диски 600GB 10k 2.5";*
- *появился SmartCache для серверов с поддержкой систем хранения 3PAR;*
- *появилась технология назначения приоритетов в обслуживании для хостов (QoS);*
- *объявлена поддержка Open VMS, Ubuntu;*
- *объявлена поддержка интерфейсов и драйверов OpenStack. Массив 3PAR еще больше встраивается в "облака";*
- *теперь подключение по FCoE поддерживается на устройствах с несколькими hopt'ами.*

*Андрей Вересов,
HP в России.*

HP расширяет портфель решений Converged Storage

Июль 2013 г. — HP расширила и обновила свой портфель решений Converged Storage. Из наиболее значительных анонсов: во-первых, была анонсирована доступность массива — HP 3PAR StoreServ 7450, построенного полностью на флэш-накопителях (SSD-накопителях). Во-вторых, HP объявила новый StoreOnce Virtual Storage Appliance (VSA), позволяющий снизить стоимость бэкапирования небольших сайтов до 65% (по результатам внутренних исследований HP).

HP 3PAR StoreServ 7450 Storage system поддерживает производительность более чем 550 000 IOPs с задержками менее 0,7 мс (по результатам внутренних исследований HP). HP 3PAR StoreServ 7450 имеет аппаратно поддерживаемое сжатие данных, позволяющее снизить требу-

емую емкость хранения до 50% без снижения производительности.

Одновременно HP анонсировала новую HP StoreFabric 16-gigabit (Gb) SAN-инфраструктуру, которая устраняет ее узкие места полосы пропускания, обеспечивая достижение максимальной производительности от использования флэш-массивов.

HP StoreOnce VSA разворачивается как виртуальная машина на существующих промышленных стандартных серверах. Новое решение — HP StoreOnce VSA — соединяет программно-определяемую сеть, сервер и СХД в единое целое, снижая стоимость и сложность.

Анонсирована новая HP StoreEver MSL6480 ленточная библиотека enterprise-класса с midrange-стоимостью, которая позволяет осуществлять масштабирование по производительности до 60,4 Тбайт/час и по емкости — до 3,5 Пбайт в рамках одной библиотеки.

Весь портфель решений HP Converged Storage интегрируется с новым HP Data Protector 8 backup and recovery ПО. Пользователи теперь могут защищать свои среды с числом до 1 триллиона файлов.

Стоимость и доступность

- HP 3PAR StoreServ 7450 Storage systems доступна по стартовой цене от \$99 000.
- HP 3PAR Priority Optimization Software доступно с июля.
- HP StoreOnce VSA доступно с июля по цене \$3 500 за 10 Тбайт емкости.
- Ленточная библиотека HP StoreEver MSL6480 Tape Library доступна по цене от \$30 000.
- HP Data Protector 8 доступен по стартовой цене от \$1 238.

Уже в августе 2013 г. HP сделала апгрэйд VSA, а также представила ряд новых решений.

В составе обновленной программно-определяемой (software-defined) СХД HP StoreVirtual Virtual Storage Appliance пользователям стал доступен ряд новых возможностей, в частности, технология automated storage tiering, работающая на уровне подтомов, а также улучшенная интеграция с VMware и Microsoft. Кроме того, HP представила новые модели лицензирования, позволяющие пользователям приобрести только тот функционал, который ему нужен сейчас, и докупать нужное по мере необходимости.

Наконец, HP представила два программных плагина для СХД HP 3PAR StoreServ Storage (один — для VMware, другой — для SAP®), которые упрощают управление сложными средами приложений.

Гибкая автоматизированная программно-определяемая архитектура хранения

По всему миру организации разного масштаба, от небольших отделений и филиалов до штаб-квартир крупных корпораций, а также поставщики облачных сервисов, все чаще выбирают программно определяемые СХД, которые не зависят от аппаратного обеспечения и гипервизора. Эта технология позволяет исполь-