

HPE: будущее post-NAND СХД

Обсуждение этапов развития следующего поколения массивов HPE 3PAR на период 2018–2023 гг., прежде всего, в контексте использования протокола NVMe over Fabrics и новых технологий для накопителей. Первый этап — интеграция новых флеш-технологий в существующие массивы, второй — постепенная интеграция NVMeOF с использованием новых ASIC Gen6 уже в новых массивах 3PAR. Публикация подготовлена на основе материалов HPE — <http://hpe.to/60078K5ez> и http://www.theregister.co.uk/2016/07/01/storeserv_asic_architecture/.

От SCSI к NVMe и новым флеш-технологиям¹⁾

Необходимость нового протокола взамен SCSI, который разрабатывался для жестких дисков много десятилетий тому назад, стала очевидна после того, как использование флеш-памяти NAND позволило резко повысить скорость работы бизнес-приложений. Оказалось, что потенциал SSD невозможно раскрыть полностью — ориентированный на «механические» носители протокол SCSI стал узким местом. Возможность одновременного выполнения множества параллельных операций (до 64К; до 64К команд в каждой операции, *прим. ред.*) в случае работы с флеш-накопителями сильно отличается от ситуации, когда данные могут записываться/считываться через фиксированное число головок (в пределах 10; при этом фактически и физически одновременно поддерживается только один поток данных для одной операции в/в, *прим. ред.*) жестко закрепленных на одном коромысле. Итак, появился новый протокол — Non Volatile Memory express (NVMe), лишенный «исторических проблем» SCSI (например, в SCSI поддерживается нескольких сотен команд, в основном, для обеспечения «обратной» совместимости; в NVMe — только 13 основных/обязательных команд, соответственно, — дополнительное снижение задержек, *прим. ред.*). Считается, что этот протокол поможет резко уменьшить время доступа и увеличить пропускную способность для бизнес-приложений.

Первыми от использования NVMe выигрывают серверы. Это неудивительно — флеш тоже сначала активнее использовался в серверах и только через какое-то время пришел в системы хранения. Однако внедрение NVMe в массивы all-flash — задача гораздо более сложная. Этот протокол может внедряться на 3-х уровнях:

- при передаче данных от серверов к СХД;
- внутри контроллеров СХД;
- при передаче данных от контроллеров СХД к накопителям.

«В настоящий момент передача данных от серверов к СХД осуществляется преимущественно по Fibre Channel. Поддержка FC-NVMe уже реализована в 6-м поколении коммутаторов Brocade/HPE, появляются соответствующие серверные HBA,

заканчивается работа над стандартизацией. Альтернативные способы применения NVMe для SAN гораздо сложнее, так как потребуют полной перестройки сети хранения в организации», — отмечает Тоня Филоненко (*Tonya Filonenko, Social Media Lead for Hewlett Packard Enterprise Russia*).

Внутри систем хранения в настоящее время используется SAS — высокопроизводительный протокол, хотя и основанный на SCSI, но доказавший за многие годы свою уникальную способность справляться с сотнями накопителей. Парадоксально, но, похоже, что внедрение NVMe для передачи данных от контроллеров до накопителей не станет актуальным до массового появления еще более быстрых устройств хранения, основанных на других физических принципах (начало поставок — 2 кв. 2017 г., *прим. ред.*). А вот внутри самого контроллера все гораздо проще. Так же, как и когда-то в начале использования флеш-памяти в серверах, можно установить ускорители и не менять сразу всю архитектуру СХД и заодно SAN.

Многие вендоры разрабатывают технологии хранения, которые будут в 10 раз быстрее существующих флеш-накопителей. По скорости доступа они еще сильнее приблизятся к оперативной памяти, но при этом будут сохранять данные при отключении питания, и будут гораздо дешевле. Общее название для этих технологий — Storage Class Memory (SCM). Сюда относятся Magnetoresistive Random-Access Memory (MRAM), Resistive RAM (ReRAM), Phase Change Memory (PCM), мемристоры HPE и Intel 3DXPoint. Приближаются к коммерческой реализации проекты Samsung Z-SSD и Western Digital 3D ReRAM, но первым на рынок, видимо, выйдет Intel (*план выхода продуктов Intel SSD Optane и их отгрузки объявлен 15 марта 2017 г., прим. ред.*).

В декабре 2016 г. на конференции Discover в Лондоне (<https://www.hpe.com/events/discover/>) компания Hewlett Packard Enterprise представила технологию «HPE 3PAR 3D Cache». 14 марта 2017 г. состоялось первое представление преимуществ данной технологии с использованием новых флеш-карт — Intel® Optane™ SSD DC P4800X Series (поставляются с технологией Intel® Memory Drive Technology, обеспечивающей прозрачную интеграцию Intel® Optane™ SSD DC P4800X в единый пул памяти без каких-либо изменений в ОС или/и приложениях).

Алгоритмы, отработанные за последние годы для NAND, прекрасно работают и с «ускорителями» Intel 3D XPoint SCM. Результаты тестов показывают снижение времени доступа на 50% и увеличение количества операций ввода-вывода на 80%.

Для тестов использовалась двухконтроллерная модель 3PAR 20450 AFA с восемью накопителями 15.36TB SSD. В каждый контроллер был дополнительно установлен Intel Optane для реализации функционала 3D Cache.

От ASIC Gen 5 к ASIC Gen 6 и NVMe over Fabrics

Одним из ключевых элементов массивов HPE 3PAR являются специализированные чипы — ASIC, которые разрабатываются с расчетом на пятилетний срок эксплуатации массива (одновременно он определяет и соответствующее поколение ASIC и массива). Существующие массивы 3Par построены на ASIC 5-го поколения и проектировались для периода с 2012 г. по 2017 г. Проектирование 6-го поколения ASIC ориентировано уже на период с 2018 г. по 2023 г.

5-е поколение ASIC разрабатывалось для построения СХД от массивов с использованием «чистых» HDD и гибридных (HDD+SSD) накопителей до AFA-массивов. Сиамак Назари (*Siamak Nazari*), архитектор HPE StoreServ и один из разработчиков ASIC Gen 6 для 3PAR, считает, что высокопроизводительные флеш-накопители (типа 3D NAND) или post-NAND накопители следующего поколения будут строиться с использованием технологий: Resistive RAM (ReRAM), Memristor, 3D XPoint (продвигаются Intel и Micron, *прим. ред.*), STT-RAM (Spin Transfer Torque RAM) и PCM (Phase Change Memory). Соответственно, ASIC Gen 6 должен учитывать особенности этих технологий, например, низкую задержку и, возможно, замену NAND-блока на байт при «сборе мусора» и др.

Назари отмечает особое значение для HPE технологий XPoint и ReRAM. При этом партнерство HPE-SanDisk (теперь WDC) продолжается и развивается на основе требований, предъявляемых к серверам HPE.

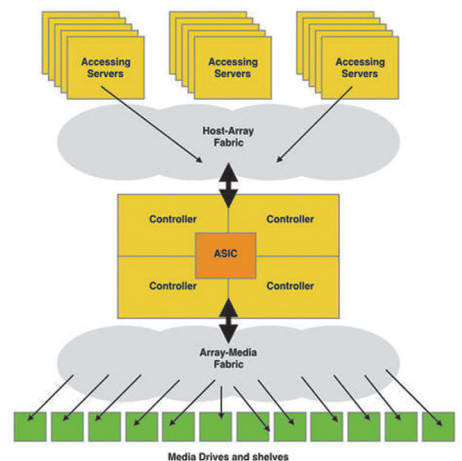


Рис. 1. Структурная схема массивов StoreServ.

1) Источники:

— HPE 3PAR 3D Cache: Turbocharged Application Acceleration Is Here (<https://community.hpe.com/t5/Around-the-Storage-Block/HPE-3PAR-3D-Cache-Turbocharged-Application-Acceleration-Is-Here/ba-p/6922573#.WEgeD-YrkKh>);
— 3PAR and Storage Class Memory: From Ludicrous Speed to Plaid (<https://community.hpe.com/t5/Around-the-Storage-Block/3PAR-and-Storage-Class-Memory-From-Ludicrous-Speed-to-Plaid/ba-p/6900913#.WEgeYUrkKh>);
— NVMe and the Need for Ludicrous Speed (<https://community.hpe.com/t5/Around-the-Storage-Block/NVMe-and-the-Need-for-Ludicrous-Speed/ba-p/6887059#.WEgeYUrkKh>).

Назари видит использование носителей ReRAM и XPoint (с их субмикросекундной задержкой доступа) в составе серверов HPE в форме DIMM (появление Intel 3D XPoint DIMM ожидается в конце 2017 г., *прим. ред.*). Использование NVMe over Fabrics ожидается для организации сетевых хранилищ нового поколения.

Назари выделяет следующие шесть основных компонент для массивов StoreServ: доступ к серверным хостам (HBA), сетевая фабрика (как правило, Fibre Channel), комплекс контроллеров массива с ASIC, внутренняя фабрика для контроллеров и накопителя массива (рис. 1).

Назари считает, что в настоящее время внутренняя фабрика для контроллеров будет строиться на Fibre Channel и SAS, а в дальнейшем будет мигрировать на архитектуру NVMe fabric (в течение двух–трех лет). ASIC во многом будет дорабатываться и будет иметь встроенную семантику хранения в предположении единого системного домена. XPoint SSD будет иметь задержку примерно 10 мкс по сравнению с 60–100 мкс в существующих NVMe NAND. SAS добавит к этому 10–20 мкс и будет иметь четко определенные протоколы обработки ошибок, в которых отсутствует NVMe. По его мнению, латентность SAS “просто прекрасна”, а протокол более надежный, чем NVMe.

Назари заявляет, что ASIC Gen 6 будет оптимизирован для работы с NVMe over fabric. При этом задержки для таких сервисов данных, как моментальные снимки и репликация, будут соответствовать эпохе Gen 6 ASIC, то есть снижены. Также он отмечает, что массивы ZPAR обычно используют половину циклов ЦП или ядер, необходимых для конкурирующих массивов, потому что ASIC с пропускной способностью 800 Мбайт/с функционирует подобно сопроцессору памяти (при этом ядра Intel Xeon характеризуются им как имеющие пропускную способность 150 Мбайт/с).

Греф сказал — в ЦРТ уже сделали

Март 2017 г. — В ходе лекции в Московском физико-техническом институте председатель правления Сбербанка Герман Греф отметил, что одной из перспективных технологий идентификации клиентов банка являются системы, не только распознающие голос, но и отслеживающие движения губ.

Такие системы уже давно привлекают разработчиков. Один из примеров — комбинация голосовой и лицевой биометрии в продуктах линейки VoiceKey от ЦРТ. Решение российского разработчика предлагает пользователю посмотреть в камеру устройства (смартфона или компьютера) и произнести случайную комбинацию цифр, выведенную на экран.

Голосовая биометрия сочетает в себе проверку не только физических характеристик человека, но и поведенческих — его манеры говорить и интонации. К ней добавляется биометрия по лицу — сравнение лица смотрящего в камеру с фото пользователя из базы. Также система проверяет соответствие движения лицевых мышц про-

гносимым словам. Минимальная рассинхронизация звучащей речи и движения лица не позволит пройти проверку. Так работает механизм liveness detection (детекция живого пользователя). Дополнительная защита обеспечивается динамической (случайной) парольной фразой, что не позволяет использовать предзаписанные аудио-или видеозаписи с целью обмана системы.

Такое сочетание распознавания речи и синхронных движений губами очень сложно для подделки. Здесь проверяются не две отдельно взятые биометрические модальности, а сочетание множества признаков. Одна только биометрическая модель голоса строится по 74 уникальным характеристикам. Это позволяет в разы повысить уровень безопасности и приблизить точность распознавания личности пользователя к 100%.

Удобство сочетания лица и голоса также заключается в их доступности для банка и клиента: камера и микрофон сегодня есть в любом смартфоне, никакого дополнительного оборудования, вроде сканера радужки и отпечатка пальца, не требуется.

«Информация, которую хранят банки, зачастую уже сама представляет собой реальные деньги, и её безопасность затрагивает интересы не только самой организации, но и очень большого числа клиентов — физических и юридических лиц. В максимальной защите нуждаются, прежде всего, информация по транзакциям и персональные данные, — комментирует генеральный директор ЦРТ Дмитрий Дырмовский. — При этом современным банкам приходится балансировать на грани безопасности и открытости — при максимальной безопасности данных нужно предоставить клиенту быстрый и удобный доступ к ним».

Необходимость наличия надежных инструментов идентификации личности, а также стремление минимизировать риски потери информации и денежных средств клиентов, являются одной из важнейших и актуальных задач для компаний, функционирующих в цифровую эпоху. В связи с этим биометрические технологии вызывают большой интерес.

Такие технологии уже успешно применяются в зарубежных и активно тестируются в ряде российских банков, например, мобильное приложение VoiceKey.OnePass от ЦРТ сейчас тестируется в ВТБ24.

IBM запускает ML в частном облаке

Март 2017 г. — Компания IBM представила IBM Machine Learning (ML), первую когнитивную платформу для непрерывного создания, обучения и развертывания большого объема аналитических моделей в частном облаке, которое лежит в основе обширных корпоративных хранилищ данных. Даже используя самые современные методы, специалисты по обработке данных, которых сейчас не хватает на рынке, могут потратить дни или недели на пошаговую разработку, тестирование и модификацию все одной аналитической модели.

IBM взяла за основу технологию машинного обучения платформы IBM Watson и,

прежде всего, делает ее доступной там, где размещена большая часть корпоративных данных заказчиков: на мейнфреймах z Systems, операционных ядрах глобальных организаций. С их помощью банки, предприятия розничной торговли, страховые, транспортные и государственные компании ежедневно проводят миллиарды транзакций.

IBM Machine Learning позволяет специалистам по обработке данных автоматизировать создание, обучение и развертывание операционных аналитических моделей, поддерживающих:

- любой язык (например, Scala, Java, Python);
- любой популярный фреймворк для машинного обучения (например, Apache SparkML, TensorFlow, H2O);
- любой тип данных по транзакциям;
- перемещение данных в облако без дополнительных расходов, задержек или рисков.

Cognitive Automation for Data Scientists, разработанная IBM Research, помогает специалистам по обработке информации выбирать подходящий алгоритм для анализа путем сравнения доступных алгоритмов с имеющимися данными и их ранжирования. Таким образом, система находит наилучшее соответствие для текущих потребностей. Сервис также учитывает различные обстоятельства, например, необходимый функционал алгоритма и скорость получения результатов.

IBM Machine Learning сначала будет доступна на z/OS, а затем появится на других платформах, включая IBM POWER Systems.

Qrator и Wallarm: отчет по сетевой безопасности

Февраль 2017 г. — Компании Qrator Labs и Wallarm представили отчет о состоянии сетевой безопасности, в котором отражены основные тенденции и события 2016 года.

Масштабность атак

Инциденты, связанные с атаками типа «отказ в обслуживании» вновь на слуху — но теперь грамотно выполненные атаки уже угрожают доступности целых регионов. На проблему вновь нужно обращать повышенное внимание — словно мы вернулись на 5–7 лет назад в прошлое.

Мощность атак и их сложность в минувшем году выросли радикально. В прошлом даже мощные атаки в 100–300 Гбит/с не вызывали особой «головной боли». Сложные типы атак на протоколы прикладного уровня случались редко. В 2016 году мир впервые увидел атаки в 1 Тбит/с, и атаки на уровень L7 стали куда более распространёнными.

Упрощение атак

Все эти годы эволюция информационных технологий шла по своего рода «пути наименьшего сопротивления». Компании боролись против времени и конкурентов, а победителем выходил тот, кто удачно и вовремя

(продолжение — стр. 31)