

# Системы хранения медиаданных сегодня и завтра

*Сергей Платонов,*

*руководитель исследовательской лаборатории RAIDIX*

**М**ы живем в мире, по-прежнему подчиняющемся закону Мура: вычислительная мощность современных информационных систем удваивается каждые 18 месяцев. Основным двигателем развития технологий является реакция рынка на постоянно растущие запросы пользователей. Не становится исключением рынок СМИ и индустрии развлечений.

В отличие от скорости развития вычислительных компонентов прогресс систем хранения данных (СХД) не столь стремителен. Линейная производительность жестких дисков выросла за десять лет менее чем в два раза. Конечно, время отклика значительно улучшились за счет использования буфера большого объема и оптимизаций операций чтения/записи.

Видео с разрешающей способностью 4К, ставшей на сегодняшний день реальностью, требует в четыре раза большей пропускной способности системы хранения данных по сравнению с 2К. Сейчас встречаются проекты, ориентированные на работу с видео 5К, 6К и 8К.

Стереоизображение 3D удваивает и эти запросы, при этом для достижения хорошего зрительного эффекта и более плавного воспроизведения видео производителям приходится увеличивать частоту кадров до 48 и даже 60 Гц. Доступные сегодня камеры, используемые при съемке спортивных событий, поддерживают скорость съемки 120 кадр/с и даже выше.

Расширенный динамический диапазон (HDR) имеют изображения с глубиной цвета 16 бит на канал. В итоге получается более чем 20-кратный рост требований к пропускной способности сети хранения данных: видеопоток 4К 3D-стерео, имеющий глубину цвета 16 бит на канал и повышенную частоту кадров, нуждается в пропускной способности 5400 МБ/с. Сравните

этот показатель с 200 МБ/с, которые необходимы для традиционного формата стандартной четкости.

Стремительными темпами увеличиваются и требования к объемам систем хранения. В проектах, где применяется технология Free Viewpoint, задействуется до 32 камер, снимающих видео 4К и генерирующих огромный поток данных.

Многие студии не могут себе позволить квалифицированный персонал, способный управляться со сложной инфраструктурой хранения, а потому они желают видеть системы, работающие по принципу plug-n-play. Чем же могут ответить производители систем хранения данных на такой рост требований?

## Интерфейсы

Наиболее распространенными в медиасфере блочными интерфейсами являются Fibre Channel, iSCSI, SAS и PCI-E. В данной статье рассматривается развитие интерфейсов Thunderbolt и USB, пригодных для небольших DAS.

Производительность Fibre Channel удваивается каждые 3 года. Официально интерфейс 16 Гбит/с появился в 2011 году, но первые устройства для медиарынка стали доступны только во второй половине прошлого года. Одними из первых поддержку нового интерфейса обеспечила компания RAIDIX, производящая программное обеспечение, способное превращать стандартное серверное оборудование в систему хранения данных высокой производительности.

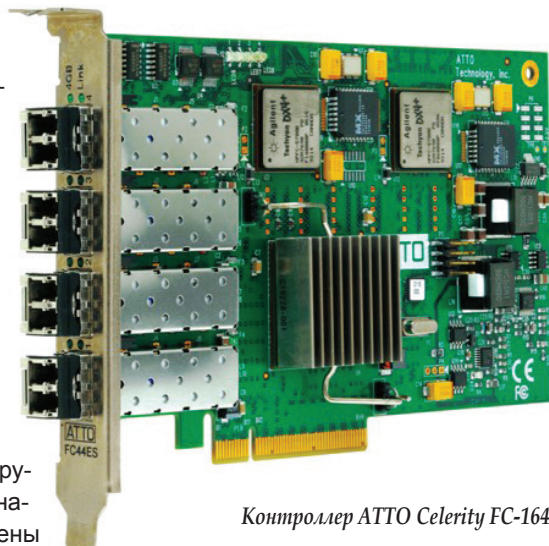
Далеко не все студии спешат переходить на новое оборудование, и связано это со значительным превышением цены

16-гигабитных коммутаторов над 8-гигабитным сетевым оборудованием. Для небольших студий решением этой проблемы может стать использование СХД, обладающей широким набором интерфейсов.

СХД, построенные с использованием программного обеспечения RAIDIX, можно оснастить более чем 16 портами Fibre Channel 16 Гбит/с с использованием контроллеров ATTO Celerity FC-164E.

Следующее увеличение производительности Fibre Channel ожидается в 2014 году. Уже не раз высказывались предположения, что рынок FC-устройств пойдет на убыль, но альтернативные технологии не обладают достаточной зрелостью и доверием специалистов, чтобы вытеснить проверенные решения с рынка, и наверняка в 2015 году появится множество устройств, оснащенных интерфейсом Fibre Channel 32 Гбит/с.

С выпуском процессоров на архитектуре Sandy Bridge появились платформы, снабженные IO-шиной PCI-E 3.0. Спецификации шины были утверждены в ноябре 2010 года, а первую материнскую плату компания Gigabyte представила в 2011 году. Не-



Контроллер ATTO Celerity FC-164E

смотря на то, что скорость передачи данных была увеличена с 5 до 8 ГТ/с (гигатрансфер в секунду), пропускная способность шины увеличилась вдвое – за счет изменения механизма кодирования. Теперь производительность одной линии достигает 985 МБ/с в каждую сторону.



Разъем mini-sas HD

Вскоре должны появиться устройства с QSFP-интерфейсами, обладающие пропускной способностью до 3900 МБ/с. Увеличение пропускной способности PCI-E дало толчок развитию других технологий: сегодня можно видеть 16-портовые SAS-адаптеры, четырехпортовые Fibre Channel HBA на 16 Гбит/с и двупортовые Ethernet NIC на 40 Гбит/с. Следующее поколение шины ожидается в 2016 году.

Во второй половине 2012 года были представлены первые устройства SAS 12 Гбит/с, созданные в развитие технологии SAS. Первым шагом внедрения новых стандартов стало появление интерфейсных разъемов SFF-8643 и SFF-8644, называемых mini-sas HD. Новые разъемы позволили значительно увеличить плотность размещения портов на HBA, поддерживающих SAS 2.1, они станут основными разъемами для устройств SAS 3.0.

За счет использования специализированных буферов на платах интерфейс SAS 3.0 сможет достигать полной производительности даже при использовании накопителей производительностью 6 Гбит/с. SAS – это не только интерфейс подключения накопителей к шине СХД, но и очень перспективный хост-интерфейс. Высокая пропускная способность широкого порта 48 Гб/с в сочетании с 4-портовыми HBA, низкий уровень латентности, доступность коммутаторов и оптических кабелей длиной до 100 м (20 м для медных кабелей) делают его прекрасной альтернативой Fibre Channel при значительном снижении цены решения.

Поддержка интерфейса SAS в скором времени появится в программном обеспечении компании RAIDIX, что сделает его самым универсальным на рынке, поддерживающим рекордное количество протоколов и интерфейсов.



Адаптер FastFrame

Все чаще можно слышать, что Ethernet в ближайшее время станет основным транспортом центров обработки данных. Оборудование Lossless Ethernet позволяет выполнять передачу данных в сетях без потерь и с обеспечением уровня обслуживания для каждого класса трафика. Часто Lossless Ethernet и используемую для него технологию Data Center Bridging ассоциируют с протоколом Fibre Channel over Ethernet. Хотя DCB может применяться и для iSCSI, упрощая его использование, позволяя обеспечивать лучшее качество сервиса (QoS) и повышая его производительность. При этом iSCSI не требует от оборудования поддержки DCB, оставляя ее как опцию, в то время как для FCoE поддержка Data Center Bridging обязательна.

Системы Lossless Ethernet построены на базе Ethernet-оборудования 10 Гбит/с с поддержкой DCB. NIC 40 Гбит/с представляют собой объединенные в один порт четыре 10-гигабитные линии. Для Ethernet 10 Гбит/с основными разъемами являются SFP+, RJ-45 и QSFP.

Адаптеры FastFrame компании ATTO – это конвергентные устройства, поддерживающие технологию DCB и используемые для подключения к FCoE, iSCSI СХД, а также к файловым системам хранения.

Большинство блочных Ethernet СХД на медиарынке поддерживает сегодня протокол iSCSI, хотя встречаются системы хранения данных FCoE и AoE.

Следующим шагом в развитии Ethernet станет появление в ближайшее время устройств производительностью 25 Гбит/с.

## Протоколы



**ПОРТАТИВНЫЕ ВИДЕОМОНИТОРЫ**

# НИЧЕГО ЛИШНЕГО




Новые портативные видеомониторы LogoVision

- ▶ ЛЕГКИЙ КОМПАКТНЫЙ КОРПУС
- ▶ ПОДДЕРЖКА HD-ФОРМАТА
- ▶ ИЗМЕРИТЕЛЬНЫЕ И СЛУЖЕБНЫЕ ФУНКЦИИ
- ▶ ПРИВЛЕКАТЕЛЬНАЯ ЦЕНА

[www.proland.ru](http://www.proland.ru)

Компания Microsoft одновременно с анонсом новых операционных систем сообщила о выходе протокола SMB 3.0. Основными новшествами в протоколе стали значительно повышенная производительность с использованием технологий SMB Direct и SMB Multichannel, а также прозрачная отработка отказа и защита от отказов в работе узла кластера с помощью SMB.

Новые важные возможности SMB 3.0:

- ◆ поддержка дистанционного снимка VSS;
- ◆ многоканальность для масштабирования производительности сети;
- ◆ SMBDirect для максимального использования производительности с помощью RDMA;
- ◆ прозрачный переход на резерв для повышения надежности;
- ◆ шифрование Plus, заимствование директорий, BranchCache v2 и др.

Во время технических сессий Microsoft показала, что теперь файловый сервер способен выполнять передачу информации с производительностью в несколько ГБ/с.

Системы на базе SMB 3.0 смогут в будущем в некоторых случаях заменить блочную СХД, так как они гибче и не требуют использования дорогостоящих кластерных файловых систем. Осталось только дожидаться клиентских приложений для отличных от Windows операционных систем.

Развитие облачных технологий и рост количества мобильных устройств приведет к тому, что файловые системы хранения будут вытеснены из привычной для них ниши. Сейчас многие аналитики называют объектные системы хранения данных NAS 2.0.

## Шасси

Требования к высокому уровню производительности заставляют разработчиков решений искать новые подходы в конструировании многодисковых систем хранения данных. Они должны быть компактны, не иметь «узкого места» (единой точки отказа) и обеспечивать должный уровень охлаждения.

Компания RAIDIX проводит в своей лаборатории тестирование шасси, способных вместить в корпусе высотой 4U 60 и 80 жестких 3,5" дисков.

Подобная плотность достигается за счет вертикальной загрузки жестких дисков в шасси. При этом обеспечивается горячая замена накопителей, блоков питания и IO-модулей. На каждом IO-модуле располагается 4 широких порта SAS 2.1, которые могут обеспечить суммарную для всех дисков пропускную способность. Модули работают в режиме «активный – активный» и позволяют выполнять зонирование дисков на порты для большей гибкости системы.

## Накопители

Механическим жестким дискам предрекают «смерть» в течении пяти

лет, однако если верить последним прогнозам, данная технология будет использоваться по крайней мере до 2025 года. Но, несомненно, использование HDD будет создавать все больше и больше проблем. Связано это с несоответствием изменения объемов и производительности накопителей. Также нет и кардинального роста надежности жестких дисков.

Уже сейчас многие специалисты говорят о том, что уровня RAID 6, способного работать при отказе двух жестких дисков, недостаточно для современных систем хранения, ведь время восстановления замененных дисков может занять несколько суток. Технологии компании RAIDIX позволили ей вывести на медиарынок уникальную функцию – массив RAID N+3, сохраняющий работоспособность при отказе 3 жестких дисков в одной группе, обнаруживать и исправлять так называемые «тихие повреждения данных» на дисках SATA, для которых, как для SAS, неприменимы технологии отслеживания ошибок T10 PI. При этом по производительности подобные массивы не уступают, а подчас и превосходят массивы RAID 6 конкурентных систем. В сочетании с технологиями упреждающего восстановления и частичной реконструкции, заметно снижающей время восстановления дисков в массивах, и гибкими механизмами буферизации RAIDIX позволяет строить оптимальные решения для требовательных к производительности и надежности. Ожидается, что на смену технологии RAID, созданной в 1980-х годах, в ближайшие четыре года придут другие решения, позволяющие обеспечить работоспособность при множестве одновременных отказов накопителей и исправлять ошибки за минуты.

Альтернативой жестким дискам являются накопители на flash-памяти. Основными тенденциями развития flash-накопителей являются:

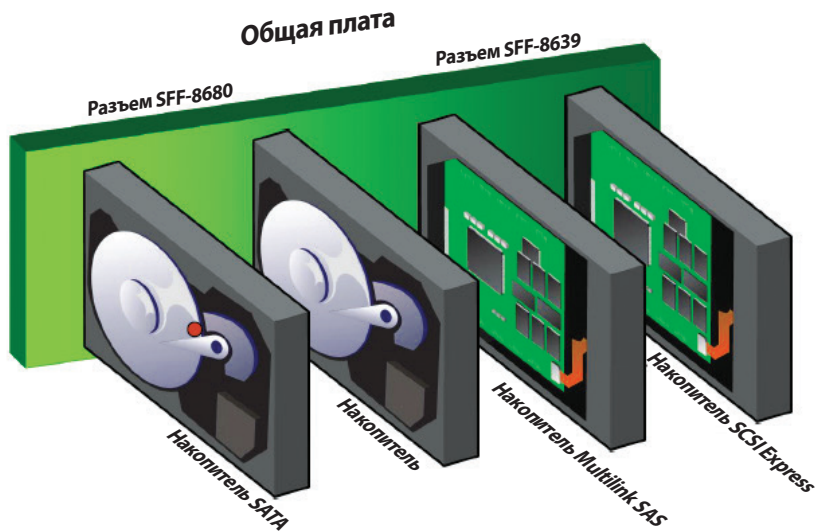
- ◆ повышение плотности чипов памяти;
- ◆ обеспечение целостности данных и долговечности микросхем;
- ◆ обеспечение гарантированной производительности.

Ожидается, что в течении двух лет объем flash-накопителей увеличится вчетверо, а показатель удельной стоимости (ГБ/\$) снизится в 3 раза.

В 2012 году стал очевиден переход к использованию устройств на памяти



Шасси с вертикальной загрузкой дисков



Общая плата с подключенными к ней различными накопителями

типа MLC. Производители приближают надежность и производительность данного типа памяти к показателям SLC. Добиваются они этого за счет усовершенствований контроллеров памяти, улучшая алгоритмы записи, оптимизируя механизмы отслеживания и коррекции ошибок, делая их незаметными для основных задач.

Сейчас много говорят о появлении TLC-памяти, но данная технология еще не будет готова к появлению на корпоративном рынке в ближайшие два года. В отчетах лабораторий нередко упоминаются успехи в адаптации PCM, STTMRAM и RRAM. Но реальных устройств, поддерживающих данные технологии, стоит ожидать не ранее 2016 года.

В 2013 году появятся первые устройства, в которых используются новые интерфейсы: SCSIExpress, NVMe и SATAExpress. NVMe будет обеспечивать уровень задержек в 3

раза меньший, чем SAS/SATA. Новый универсальный разъем SFF 8639 позволит использовать устройства, подключаемые к общей плате системы, с возможностью горячей замены.

При этом общая плата сможет поддерживать интерфейсы SAS 12 Гбит/с, SATA 6 Гбит/с, SATA Express, SAS Express, NVMe/SOP PQI, многоканальный SAS (4-портовые устройства).

### Архивирование

Самой ресурсоемкой с точки зрения использования пространства хранения задачей является архивирование данных.

Наиболее распространенным способом архивирования является запись данных на магнитную ленту. Linear Tape File System (LTFS), с одной стороны, является форматом данных, записываемых на магнитную ленту, а с другой – программным

интерфейсом, предоставляющим доступ к данным, хранящимся в библиотеках LTO 5, как к стандартной файловой системе, в том числе и с использованием иерархической структуры и с доступом к метаданным. Такой подход значительно упрощает работу с накопителями и расширяет набор программного обеспечения, позволяющего использовать ленточные библиотеки.

В 2012 году LTFS была названа одной из основных технологий, связанных с системами хранения. Впервые LTFS была анонсирована в 2010 году и представлена на NAB компанией IBM.

Одной из тенденций архивирования является переход на оптические накопители. Основным поставщиком оборудования, популяризирующим данную технологию, это компания Panasonic. По утверждению представителей этой компании, системы архивирования на оптических накопителях способны хранить информацию в течении 50 лет, имеют гораздо меньшее время доступа, чем ленточные накопители (примерно 200 мс) и сравнимы с ленточными накопителями по объему.

В начале этого года на выставке CES Panasonic показала роботизированный RAID-архив на дисках Blu-ray RAID. Накопители в нем представляют собой картриджи на 12 дисков, объединенные в RAID. Каждый картридж обладает следующими характеристиками:

- ◆ пропускная способность – 200 МБ/с;
- ◆ емкость – 108 ТБ для архива из 90 картриджей;
- ◆ время доступа – 60 с;
- ◆ потребление энергии в режиме простоя – 6 Вт.



реклама

# RAIDIX



## СХД для монтажных студий

- До 6ГБ/с в RAID6 на чтение и запись
- Неограниченный размер LUN
- Стабильная производительность
- Поддержка A/V приложений

Raidixstorage.com