

Hitachi AMS2000: переход на SAS-архитектуру

Обзор архитектурных особенностей и функциональных возможностей нового семейства систем хранения среднего класса, анонсированного HDS в середине октября 2008 г.

Введение

В середине октября 2008 г. корпорация Hitachi Data Systems анонсировала линейку систем хранения среднего класса нового поколения – серию Hitachi Adaptable Modular Storage (AMS) 2000. В состав нового семейства входят 3 модели: AMS 2100/2300/2500. Первые две уже доступны на рынке, поставки старшей модели ожидаются в конце 2008 г. Из наиболее существенных нововведений необходимо отметить следующие:

- существенное увеличение внутреннего параллелизма системы за счет поддержки до 32 внутренних SAS-каналов точка-точка со скоростью 3 Гбит/с, что увеличивает производительность серии AMS 2000 в сравнении с предыдущей до 4 раз;
- динамическое выравнивание нагрузки между двумя контроллерами СХД серии 2000 позволяет практически без всяких усилий вывести систему хранения на пиковый уровень производительности;
- поддержка уровня доступности 99,999% (5 мин простоя в год) на всех моделях серии отражает один из основных трендов этого года – достижение СХД среднего класса (от основных поставщиков, включая EMC, HP и др.) корпоративного уровня надежности, что дает возможность использования серии 2000 в составе высококритичных приложений;
- полный переход в серии 2000 на SAS-архитектуру (включая контроллеры, интерфейсы и диски), что позволяет получить дополнительную гибкость при совместном использовании SAS- и SATA-дисков;
- пониженное энергопотребление за счет опции, позволяющей понижать/повышать RPM дисков (и останавливать их) в зависимости от частоты обращения к ним;
- поддержка больших объемов хранения: модель 2500 может иметь до 480 дисков, что при использовании SATA-дисков 7200 rpm емкостью 1 Тбайт составляет около 0,5 Пбайт. Помимо этого, поддерживаются т.н. “MegaLUN” емкостью до 60 Тбайт, что бывает очень полезным при реализации больших проектов по консолидации данных;
- анонсирование опции “thin provisioning” (или Hitachi Dynamic Provisioning), ко-

торая будет доступна с середины 2009 г., позволяет виртуализовать пул хранения, а при использовании СХД серии 2000 в составе серверных виртуальных инфраструктур (например, на базе VMware или MS Hyper-V) – и всю IT-инфраструктуру в целом.

Архитектурные особенности серии AMS2000

Система AMS2000 состоит из двух контроллеров, расположенных на них кэш-памяти и портов ввода-вывода к серверам, а также дисковых полок с носителями. Сравнение различных моделей серии AMS2000 дано в табл. 1. Схематичное представление архитектуры старшей модели – AMS2500 – дано на рис. 1.

Компоненты каждого контроллера можно условно разделить на пять категорий:

- управляющий модуль, включающий высокопроизводительный процессор Intel Xeon и выделенную оперативную память. Данный процессор осуществляет все служебные функции, необходимые для работы системы, включая управление вводом-выводом,

Табл. 1. Технические характеристики моделей серии AMS 2000

	AMS 2100	AMS 2300	AMS 2500
Кол-во дисков	120	240	480
Типы дисков	SAS, SATA,	SAS, SATA,	SAS, SATA,
Кол-во портов FC	0-4	0-8	0-16
Кол-во портов iSCSI	0-4	0-4	0-8
Кол-во кэш-памяти	8GB	16GB	32GB
Кол-во внутр. путей к дискам	16	16	32
Кол-во томов	2048	4096	4096
Типы RAID	0, 1, 5, 6, 10	0, 1, 5, 6, 10	0, 1, 5, 6, 10
Макс IOPS	400 000	450 000	900 000
Кол-во клонов	1024	2048	2048
Кол-во снапшотов	1024	2048	2048

однако не используется для непосредственного перемещения данных и задач по вычислению четности для операций RAID. Выделенная память служит для работы микрокода (или операционной системы) СХД и его служебных функций и не хранит непосредственно данные ввода-вывода;

- RAID-процессор (или DCTL) – узкоспециализированный высокопроизводительный процессор (ASIC), выполняющий функции непосредственного перемещения данных между системами ввода-вывода и кэш-памятью и расчета четности при операциях RAID;

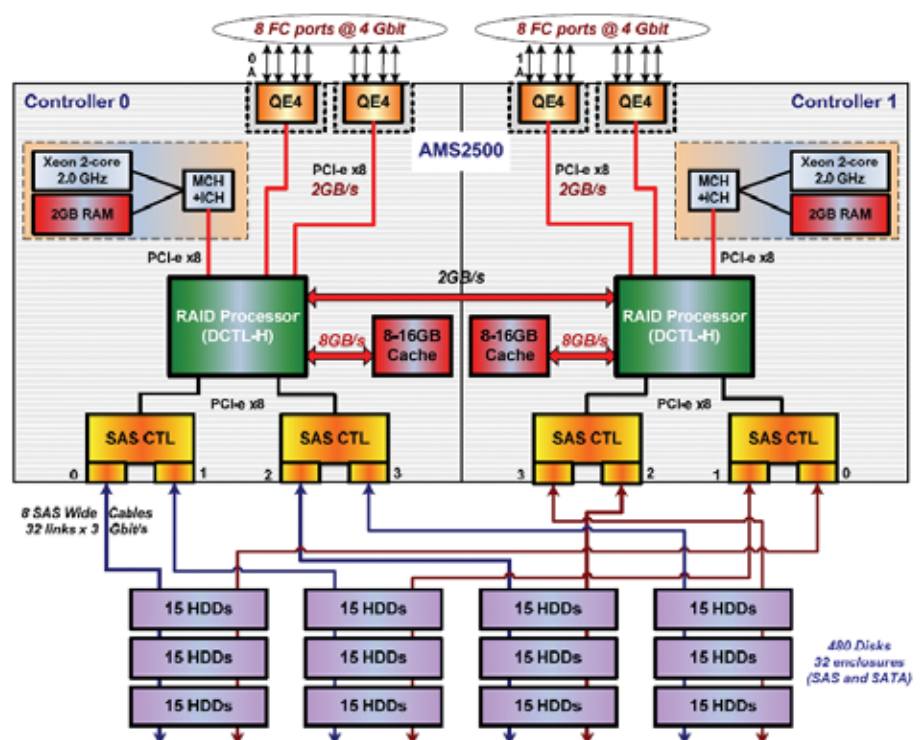


Рис. 1. Архитектура системы HDS AMS2500.

- **кэш-память**, используемая для хранения данных ввода-вывода при записи или чтении. Кэш-память служит высокоскоростным буфером между хостом и физическим диском и обеспечивает максимальную производительность и минимальное время отклика системы. Аппаратно кэш-память выполнена в виде DIMM-модулей, устанавливаемых в банки на контроллерах. В зависимости от используемых модулей и их количества, объем кэш-памяти системы может меняться;
- **модуль ввода-вывода к хостам**, состоящий из высокоскоростных процессоров Tachyon QE4, каждый из которых управляет четырьмя портами Fibre Channel скоростью 4 Gbit/sec. Аппаратно модули ввода-вывода выполнены в качестве плат расширения, устанавливаемых в специальный слот на контроллере.
- **модуль ввода-вывода к дисковой подсистеме**, состоящий из SAS контроллеров, каждый из которых обслуживает 8 активных путей 3 Gbit/sec full duplex.

Все компоненты контроллера (включая индивидуальные устройства модулей) соединены шинами PCI-Express x8 пропускной способностью 2048 MB/sec (~20 Gbit/sec). То есть, если в системе установлено 2 модуля ввода-вывода к серверам на контроллер, каждый модуль будет подключен к RAID-процессору по выделенной шине PCI-Express x8.

Исключение составляет соединение RAID-процессор <-> кэш-память, пропускная способность которого составляет 8192 MB/sec (AMS2500 и AMS2300) или 4096 MB/sec (AMS2100).

Дополнительное соединение с пропускной способностью 2048 MB/sec связывает RAID-процессоры разных контроллеров для целей зеркалирования кэш-памяти, передачи данных между контроллерами и служебных операций.

Организация работы кэш-памяти

Кэш-память в современных СХД является одним из наиболее критичных компонентов и выполняет роль, аналогичную роли оперативной памяти в серверах: предоставляет максимально быстрый доступ к данным. Через кэш-память системы происходит весь ввод-вывод от серверов – запись и чтение. При этом современные СХД используют ряд алгоритмов, позволяющих повысить эффективность использования кэш-памяти, в том числе отложенную запись (destage) на диски и предсказательное чтение (prefetch) с дисков.

Для защиты от потери данных в кэш-памяти, которые еще не были сохранены на диск, в случае сбоя используются механизмы зеркалирования содержимого кэш-памяти между контроллерами, а также техники сохранения данных при отключении питания (питание кэш-памяти от батарей в AMS2000).

Использование в AMS2000 выделенной процессорной памяти для служебных нужд, в сочетании с относительно компактным микрокодом на базе ОС реального времени VxWorks, позволяет получить значительный выигрыш с точки

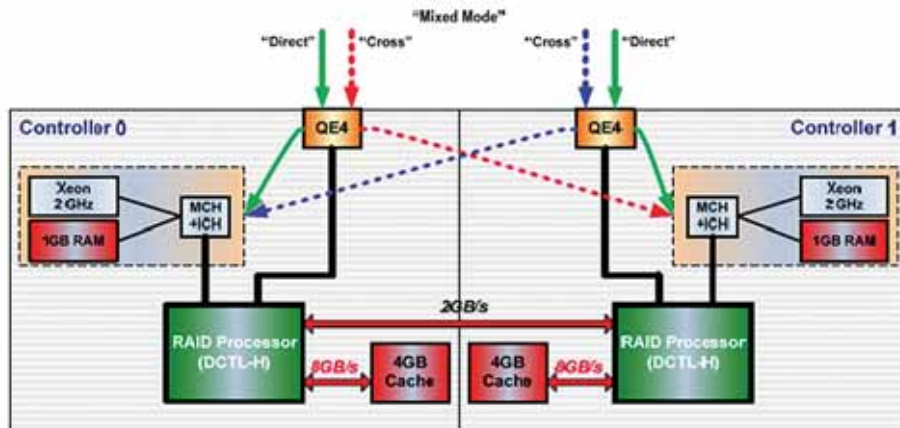


Рис. 2. Схематическое изображение технологии Symmetric Active/Active.

зрения доступной для заказчика кэш-памяти.

Система AMS2000 имеет встроенную функцию Cache Partition Manager (CPM). Данная функция позволяет выделять из пользовательской области кэш-памяти разделы и выдавать их в эксклюзивное пользование отдельным серверам или группам серверов. При этом каждый такой раздел может обладать собственными характеристиками, в частности, размером сегмента кэш-памяти, для максимального соответствия параметрам нагрузки приложения.

Подсистема ввода-вывода к серверам

Подсистема ввода-вывода в AMS2000 построена по принципу Symmetric Active/Active (рис. 2). За счет внутренней коммутации и использования отдельного процессора для перемещения данных, ввод-вывод с портов одного контроллера может быть перенаправлен на другой контроллер без использования вычислительных ресурсов первого контроллера, а, значит, без потери производительности.

Еще одной возможностью AMS2000 является автоматическая балансировка нагрузки между контроллерами (рис. 3). В случае, когда в течение определенного времени, нагрузка на контроллеры непропорциональна, система автоматически передает том другому контроллеру для выравнивания нагрузки. При этом изменения пути, по которому осуществляется доступ к тому со стороны сервера, не происходит.

В результате использования схемы Symmetric Active/Active значительно упрощается планирование системы (т.к. отсутствует необходимость ручного распределения томов по контроллерам и жесткой привязки томов к путям), улуч-

шается производительность (за счет автоматической балансировки) и устраняется необходимость в наличии специализированного ПО обеспечения доступа по множеству путей (т.е., автоматически поддерживаются любые встроенные средства ОС).

Дисковая подсистема

Дисковая подсистема AMS2000 построена на базе архитектуры SAS. Каждый контроллер системы оснащается одним или двумя (в зависимости от модели) SAS-контроллерами. К каждому SAS-контроллеру в свою очередь подключается пара интерфейсных кабелей с четырьмя путями 3 Gbit/sec Full Duplex. Пара интерфейсных кабелей (по одному с каждого контроллера) обслуживает 8 дисковых полок или 120 дисков в моделях AMS2300 и AMS2500 и 4 дисковых полки или 60 дисков в модели AMS2100. Т.о., в моделях AMS2300 и AMS2500 на каждые 8 дисковых полок приходится 8 активных двунаправленных путей 3 Gbit/sec, в модели AMS2100 – 16 активных путей на 8 полок.

Общее количество путей к дисковой подсистеме в каждой из моделей приведено в табл. 1.

Каждая дисковая полка AMS2000 оснащается парой SAS-коммутаторов, осуществляющих соединение конкретных дисков с контроллером. Выбор конкретного пути для доступа к диску осуществляется в системе автоматически с выполнением балансировки нагрузки для повышения производительности.

Дисковые полки AMS2000 являются универсальными и допускают установку дисков SAS (диски FC с интерфейсом SAS) и SATA, в том числе диски разного типа в одну полку.

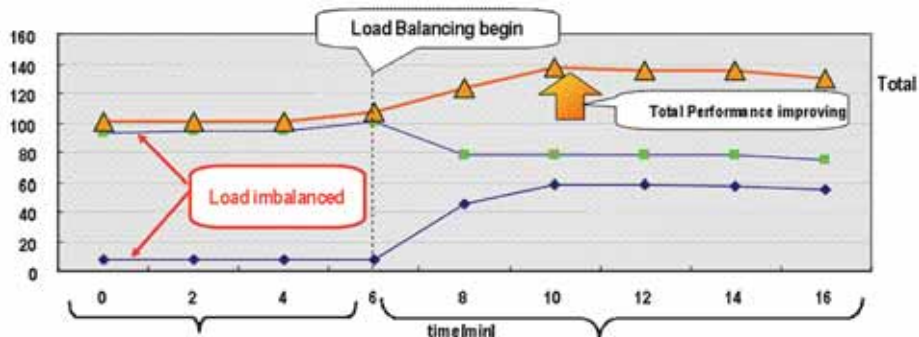


Рис. 3. Визуальное представление технологии автоматической балансировки нагрузки между контроллерами в процентах (100% – максимальная производительность контроллера).

Максимальное количество дисков в семействе AMS2000 составляет 120 дисков для AMS2100, 240 дисков для AMS2300 и 480 дисков для AMS2500.

В серии AMS2000 значительно расширено использование энергосберегающих технологий. В частности, для дисков SATA II есть возможность устанавливать 7 профилей энергопотребления. За счет этого, а также встроенных алгоритмов снижения нагрузки на диски удается экономить до 40% (по данным HDS, прим. ред.) электроэнергии.

Также на основе лицензии на отключение дисков (Power Saving) есть возможность отключения RAID-групп по команде с сервера или административного интерфейса

Динамическое расширение RAID-группы и Mega-LUN

Еще одна особенность, положительно сказывающаяся на производительности и управляемости, это возможность динамического добавления дисков к RAID-группе при этом тома автоматически перераспределяются по всей расширенной RAID-группе (рис. 4). Расширением этой опции является возможность создания Mega-LUN емкостью до 60 Тбайт при том распределяется по всем дискам системы.

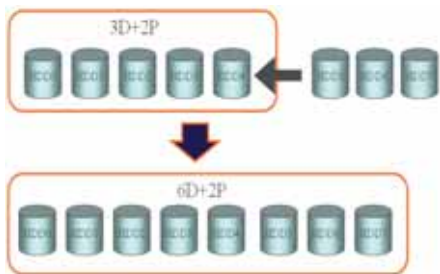


Рис. 4. Диски могут динамически добавляться к RAID-группе при этом тома автоматически перераспределяются по расширенной RAID группе.

Производительность системы AMS2000

В целом, за счет новых компонент и архитектурных особенностей по сравнению с предыдущим поколением производительность (по данным HDS, прим. ред.) контроллеров возросла до 3 раз, а производительность дисковой подсистемы — в 2-4 раза. Максимальные значения показателей производительности (по данным HDS, прим. ред.) для разных моделей серии AMS2000 даны в табл. 2.

Табл. 2. Максимальные значения производительности моделей серии AMS200

Модель/показатель	2100	2300	2500
Read Hit (IOPS)	400K	448K	910K
Random Read (IOPS)	64K	78K	130K
Sequential Read (MB/Sec)	1250	1360	2500

Заключение

Анонсирование серии HDS AMS2000 — подтверждение основных тенденций отрасли для СХД среднего класса, дающих возможность: существенно расширить архитектурную гибкость системы, а также области их применения для критических приложений; приблизить эффективность использования системы к максимальной.

Новые возможности семейства HDS USP

Октябрь 2008 г. — Корпорация Hitachi Data Systems (собственность компании Hitachi, Ltd.) объявила о доступности значительных расширений функциональности для семейства систем хранения данных Hitachi Universal Storage Platform.

Новая функция шифрования данных на уровне контроллера при записи их на диск из кэша (и дешифрование при считывании в кэш с диска) с уровнем защиты 256 бит дает возможность обеспечения конфиденциальности данных при техническом обслуживании дисков. Интегрированные средства управления ключами позволяют безопасно экспортировать ключи шифрования и хранить их резервную копию вне территории предприятия или в коммерческой службе ответственного хранения ключей, чтобы обеспечить дополнительную защиту.

HDS также представила усовершенствованные возможности программного обеспечения Hitachi Dynamic Provisioning, в том числе динамическое онлайн-расширение виртуальных томов для серверов Microsoft® Windows® и расширенную поддержку программного обеспечения Hitachi TrueCopy® Remote Replication и Hitachi Universal Replicator. Такая расширенная поддержка обеспечивает парную работу томов как с динамическим предоставлением, так и статических томов, а также онлайн-расширение объема томов с динамическим предоставлением. Улучшенное ПО Hitachi Dynamic Provisioning открывает более гибкие возможности для решения повседневных задач защиты ценных данных и выполнения процедур восстановления после аварий, при этом экономия на средствах хранения, предоставляя более простое администрирование и оптимизированную производительность.

Помимо этого, HDS ввела поддержку дисков 450 Гбайт со скоростью вращения шпинделя 15К rpm, а также две новые услуги по защите данных — уничтожение и переработку дисков (Disk Destruction and Recycling Service) и уничтожение данных на диске (Disk Data Eradication Service).

Дедупликация для систем NetApp VTL

Ноябрь 2008 г. — Компания NetApp объявила о реализации технологии дедупликации для виртуальных ленточных библиотек NetApp Virtual Tape Library (VTL), за счет чего до 95% (по результатам внутреннего тестирования NetApp, октябрь 2008, прим. ред.) уменьшается дисковая емкость, необходимая для резервного копирования на любых системах, включая не только от NetApp. Помимо этого, сокращается потребность

в ленточных накопителях и соответствующей инфраструктуре.

Технология дедупликации в составе NetApp VTL является внутренним компонентом системного программного обеспечения. Такой подход упрощает не только настройку, управление и модернизацию систем, но и снижает потребности в емкости. Технология дедупликации NetApp доступна как для действующих, так и для будущих заказчиков без дополнительной оплаты.

Среди основных преимуществ использования дедупликации в NetApp VTL называются следующие:

- 10-кратное ускорение резервного копирования при производительности и масштабируемости корпоративного класса (по данным исследования Штуттгартского университета, прим. ред.);
- 50%-ное сокращение затрат на ленточные носители за счет непосредственной интеграции с ленточными библиотеками, существенно повышающей эффективность использования ленты (по данным NetApp, прим. ред.);
- 50%-ное и более сокращение инфраструктуры ленточных библиотек за счет максимального использования ресурсов ленты путем технологии прямого создания лент в NetApp VTL (по данным NetApp, прим. ред.);
- 10-кратное ускорение восстановления данных за счет «гладкой» интеграции с ведущими приложениями резервного копирования, включая Symantec NetBackup, CommVault Simpana, BakBone NetVault, Syncsort Backup Express, и Atempo Time Navigator (по данным NetApp, прим. ред.);
- упрощение управления за счет высокомасштабируемой модульной архитектуры на базе одноплатформенного решения NetApp.

В целом, благодаря новым возможностям, системы NetApp VTL способны хранить до 10 Пбайт данных, значительно повышая как эффективность использования хранилищ, так и операционную эффективность, одновременно снижая затраты на резервное копирование.

EMC Atmos: файловое хранилище для глобально распределенных сред

Ноябрь 2008 г. — Корпорация EMC объявила о доступности своей новой разработки — EMC Atmos™ (кодовое имя разработки — Maui), ориентированной для использования в качестве файлового хранилища в глобально распределенных инфраструктурах (cloud infrastructure). EMC позиционирует данное решение как новый класс систем хранения — cloud optimized storage (COS).

Данная разработка ориентирована на интернет-провайдеров, телекоммуникационные компании и компании индуст-