

ETERNUS DX – СХД для динамических инфраструктур

В течение лета и осени 2009 г. компания Fujitsu анонсировала новую серию блочных СХД – ETERNUS Disk Storage Systems (ETERNUS DX) – от систем начального уровня до корпоративных. Одновременно были сняты с производства линейки этого семейства СХД, производимые с 2006 г. Особенностью данного анонса явилось то, что все модели нового семейства стали доступны в России и странах СНГ.



Александр Яковлев – менеджер по маркетингу продукции RISC/UNIX серверы и системы хранения, Fujitsu Technology Solutions.

Введение

Технологии, положенные в основу семейства ETERNUS, имеют многолетнюю историю и исторически активно развивались Fujitsu как системы хранения повышенного уровня надежности для собственных мейн-фреймов, суперкомпьютеров и UNIX-систем. За последнее десятилетие фокус компании стал смещаться в сторону открытых систем, и многолетние разработки систем оказались востребованными на рынке СХД для решения задач по консолидации и виртуализации. При этом до 2009 г. системы хранения Fujitsu практически не поставлялись в регион EMEA, и основные продажи этих систем приходились на Азиатский регион. В составе семейства ETERNUS есть дисковые массивы, ленточные библиотеки, виртуальные ленточные библиотеки и ряд решений для организации архивного хранения и резервного копирования с применением дедупликации. Все дисковые системы называются ETERNUS DX. В составе семейства ETERNUS DX – 8 базовых моделей (рис. 1, табл. 1). Архитектура семейства ETERNUS DX в сегменте начального и среднего уровня – классическая, в сегменте корпоративного класса используются множество контроллеров и мощная коммутационная матрица для подключения множества дисковых полок. Система DX 8100 является узкоспециализированной, и, в основ-

ном, поставляется для работы с мейн-фреймами. Все остальные дисковые системы – в полной мере системы широкого назначения и предназначены для работы с открытыми системами разного уровня. Особенности нового семейства, прежде всего, является функциональность, которая позволяет добиться максимальной эффективности в:

- производительности за счет балансировки и использования разных механизмов доступа к данным в зависимости от характера нагрузки;
- использовании емкости за счет эффективных аппаратных и программных механизмов по оптимизации;
- потреблению электроэнергии.

За счет применения отмеченных встроенных технологий модели семейства ETERNUS DX убедительно демонстрируют одни из наилучших показателей производительности как на потоковой нагрузке, так и при случайном доступе на тестах

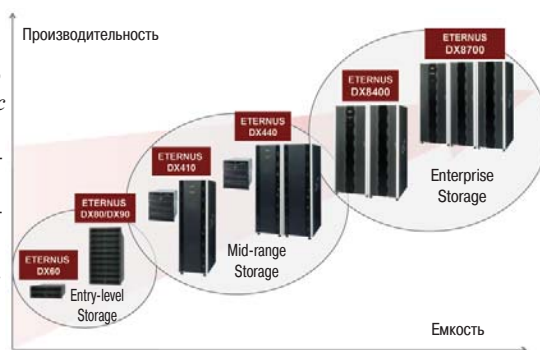


Рис. 1. Семейство моделей ETERNUS DS

SPC-1 (прежде всего это важно для таких типов критически важных приложений как Online Transaction Processing – OLTP, базы данных и почтовые серверы), а также по показателям: стоимости за единицу производительности и производительность на один диск (по результатам тестирования ETERNUS DX80 на тестах SPC-1, октябрь 2009 г.). При этом даже при максимальном количестве жестких

Табл. 1. Характеристики моделей семейства ETERNUS DS (по состоянию на апрель 2010 г.)

	DX60	DX80	DX90	DX410	DX440	DX8100	DX8400	DX8700
Class	Entry-level			Mid-range		Enterprise		
Physical storage capacity	48.0TB	240.0TB	240.0TB	414.4TB	834.4TB	118.0TB	2,008.0TB	5,456.0TB
Maximum cache capacity	2GB	4GB	4GB	8GB	32GB	16GB	256GB	512GB
Number of controllers	1 or 2	1 or 2	1 or 2	1 or 2	1 or 2	1 or 2	2 or 4	2 or 4 or 6 or 8
Number of disk drives	2 - 24	2 - 120	2 - 120	8 - 210	8 - 420	8 - 60	16 - 1,020	64 - 2,760
Макс. кол-во LUN	512	1024						
Макс. кол-во снапшотов	8-512	8-1024	8-1024	4096	4096	4096	16384	16384
Поддерживаемые диски	3.5" SSD 200GB, 100GB			400GB, 200GB				
	3.5" FC (15Krpm)			600GB, 450GB, 300GB (4Gbit/s)				
	3.5" SAS (15Krpm)			600GB, 450GB, 300GB (3Gbit/s)				
	3.5" Nearline SATA (7.2Krpm)			2TB, 1TB, 750GB, 500GB (3Gbit/s)				
	3.5" Nearline SAS (7.2Krpm)			2TB, 1TB, 750GB (3Gbit/s)				
Поддерживаемые хост-интерфейсы	FC (Max 8Gbit/s) 2 or 4 4 or 8 2 - 8 4 - 16 4 - 16 4 - 64 16 - 128							
	iSCSI (1Gbit/s) 2 or 4 2 or 4 - 2 or 4 4 or 8 4 - 8 4 - 32 4 - 64							
	SAS (3Gbit/s) 2 or 4 2 or 4 - - - - - -							
	OCLINK (Max. 17MB/s) - - - - - 4 - 8 4 - 32 4 - 64							
	FCLINK (Max. 2Gbit/s) - - - - - 4 - 8 4 - 32 4 - 64							
Поддерживаемые платформы	Windows Server™ 2003 & 2008; Solaris™ 9, 10; RedHat Enterprise Linux Server & Advanced Platform; HP-UX™; IBM AIX™; VMware vSphere™ & VMware Infrastructure™; Citrix XenServer™							
RAID levels	RAID 0, RAID 1, RAID 1+0, RAID 5, RAID 5+0, RAID 6							
Data Encryption	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Eco-mode (MAID)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

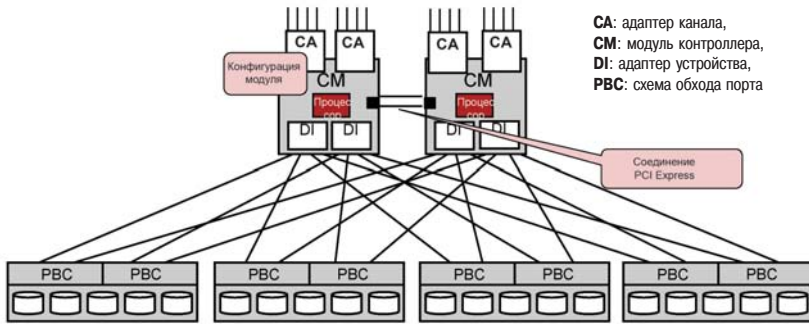


Рис. 2. Архитектура моделей семейства ETERNUS DX среднего уровня.

дисков и 100% загрузке система работает с запасом по производительности внутренних межсоединений, что позволяет легко, предсказуемо и линейно масштабировать производительность и емкость системы при апгрейдах системы во время всего срока эксплуатации.

Архитектура семейства СХД ETERNUS DX

Модели СХД семейства ETERNUS DX имеют архитектуру модульного типа (рис. 2), хорошо сбалансированную по производительности и емкости. Адаптеры канала (CA) и устройства (DI) интегрированы в модуль контроллера (CM) для увеличения скорости передачи данных.

Модели начального уровня – DX60/DX80/DX90 – комплектуются дисками SSD (Solid State Drives), SAS и Nearline SAS. Поддерживаются хост-порты: FC 4 Гбит/с, FC 8 Гбит/с (только для моделей DX80/DX90), SAS и iSCSI.

Отличительной особенностью семейства ETERNUS DS является то, что все базовые технологии и функциональность представлены во всех моделях независимо от ее класса. Отметим основные:

- функция *Data Block Guard* – добавляет 8-байтовые контрольные коды к каждому блоку из 512 байт данных, что обеспечивает максимальную целостность данных на диске, в кэше и между ними;
- функция *Redundant Copy* – позволяет заранее предсказать выход из строя жесткого диска и заблаговременно начать процедуру перемещения данных на диск горячей замены. Данная технология позволяет существенно снизить вероятность потери данных в случае выхода из строя жесткого диска;
- возможность шифрования данных на одном или нескольких логических томах аппаратными средствами дискового массива;
- множественные точки постоянного мониторинга температуры и основных параметров работы внутри дискового массива. При увеличении температуры вентиляторы системы переходят в режим удвоенной скорости работы;
- динамический диспетчер дисков (*disk traffic controller, DTC*). Кроме управления серверной подсистемой ввода/вывода, в дисковом массиве ETERNUS

CA: адаптер канала,
CM: модуль контроллера,
DI: адаптер устройства,
PBC: схема обхода порта

выполняется ряд внутренних операций, таких как восстановление данных после отказа устройства, расширенные операции копирования и т.п. Внутренние операции, как правило, выполняются последовательно, тогда как работа серверной подсистемы ввода/вывода сочетает последовательный и произвольный доступ. Для обеспечения стабильной производительности в дисковом массиве ETERNUS применяется эффективная балансировка нагрузки между внутренними операциями и серверными операциями ввода/вывода (рис. 3);

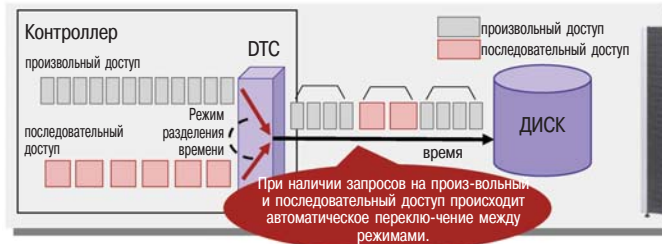


Рис. 3. Функция “динамический диспетчер дисков” (*disk traffic controller, DTC*) позволяет повысить производительность за счет оптимизации нагрузки между потоковыми и случайными операциями.

- *Eco-mode (поддерживает MAID-технологии – Massive Array of Idle Disks)* – позволяет останавливать неиспользуемые диски, за счет чего снижать энергопотребление до 20%. Может назначаться отдельным RAID-группам как по расписанию, так и по анализу запросов ввода-вывода к дисковой группе (например, когда в течение часа не происходит никакой дисковой активности). Данная технология весьма эффективна для дисковых групп, которые используются для архивного и резервного копирования данных. При этом диски запускаются автоматически при первом обращении к ним;
- *Cache Protector* – копирует кэш на флэш-память на случай сбоя электропитания в системах начального уровня. Это обеспечивает высокую доступность – после включения электропи-

тания кэш восстанавливается в течение 3-х минут и его неограниченное время хранения. Кроме того, использование конденсаторов вместо аккумуляторных батарей позволяет избежать процедуры по проактивной плановой замене аккумуляторных батарей каждые 2–3 года. Использование этой технологии позволяет существенно снизить совокупную стоимость владения системы начального уровня.

Все модели серии ETERNUS DX имеют полное резервирование, включая все компоненты – схемы контроллера, источники питания, вентиляторы и диски сконфигурированы с учетом резервирования.

Горячая замена и горячее расширение доступны для всех основных аппаратных компонентов. Возможно также обновление микрокода без прерывания работы.

В составе моделей среднего класса используются два типа процессоров: DX410 – двухъядерный процессор с тактовой частотой 2,33 ГГц (2 процессора – 4 ядра); DX440 – четырехъядерный процессор с тактовой частотой 2,83 ГГц, (2 процессора – 8 ядер).

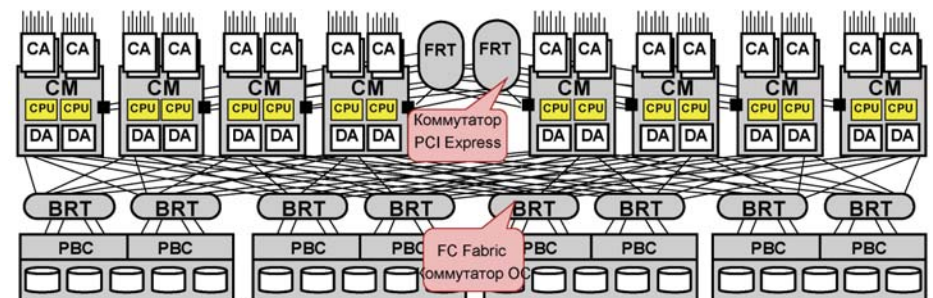
Для связи с модулем CM используются продублированные специализированные высокоскоростные шины, сделанные на базе

PCI Express и обеспечивающие избыточность и высокую производительность (см. рис. 2).

В одну дисковую полку можно одновременно устанавливать жесткие диски всех типов – FC, NL-SAS и твердотельные накопители.

Старшая серия ETERNUS –DX8000 – самая масштабируемая – до восьми контроллеров и до 2760 дисков (рис. 4). В каждый модуль контроллера ставятся по два четырехъядерных процессора с тактовой частотой 2,83 ГГц (всего до 16 процессоров или 64 ядер в системе).

Между собой модули контроллеров связываются через FRT (коммутаторы PCI Express). Подключение модулей CM и накопителей осуществляется через BRT (коммутаторы FC Fabric).



CA: адаптер каналов, CM: модуль контроллера, DA: адаптер накопителя, FRT: буферный маршрутизатор, BRT: выходной маршрутизатор, PBC: схема обхода порта

Рис. 4. Архитектура моделей семейства ETERNUS DS старшего уровня – DX8400 и DX8700.

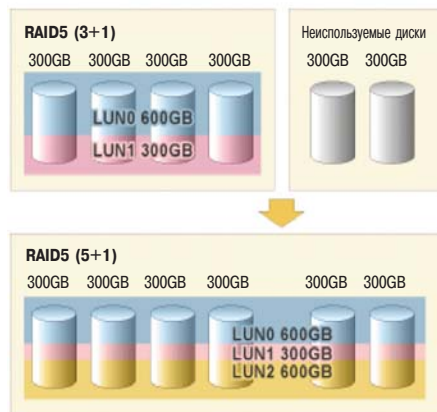


Рис. 5. Расширение RAID5 (3+1) группы до RAID5 (5+1), используя резервные диски, с определением нового LUN2.

Функциональные особенности семейства СХД ETERNUS DX

Помимо отмеченных, необходимо выделить еще ряд функциональных опций семейства ETERNUS DX, существенно повышающих эффективность использования ресурсов и упрощающих администрирование.

Онлайнное расширение RAID-групп и LUN – Logical Device Expansion

Данная опция позволяет динамически увеличивать емкость RAID-групп в реальном времени без прерывания текущих операций (рис. 5).

Динамическая миграция LUN

В процессе работы LUN могут быть перенесены:

- на другие диски;
- на другие диски и RAID-группы;
- на другие диски и диски различных типов (например, с SAS на NL SAS, с SAS на SSD).

Это позволяет сократить количество горячих точек, а также размещать LUN на более подходящих типах дисков или изменять уровень RAID в случае необходимости.

Снапшоты и клоны

Для поддержания функций резервного копирования/восстановления средства дискового массива в составе виртуализированных ЦОД в серии ETERNUS DX предусмотрены *моментальные снимки*, которые фиксируют только изменения и *клоны*, которые являются полными копиями.

Особенности первой технологии:

- меньшие требования к дисковому пространству – только для измененных данных;
- снижение производительности источника при использовании снапшота;
- более высокая латентность при первой записи на источник;
- непрерывное чтение из снапшота влияет на источник (LUN);
- потеря моментального снимка в случае удаления исходного LUN или сбоя RAID.

Данная технология рекомендована для экономных в плане занимаемого места временных копий (копирование на ленту).

Клоны являются полными копиями LUN-источника и требуют столько же места, сколько и источник. Особенности данной технологии:

- более высокая латентность при операциях чтения и записи на источнике во время процесса копирования;
- не требует дополнительных ресурсов после завершения процесса копирования;
- клон полностью независим от источника.

Данная технология рекомендована для создания долговременных копий (клоны БД, восстановление с диска).

На базе обеих технологий возможно использование клона логического тома, на который по заданному расписанию могут переноситься изменения в основном логическом томе.

В составе ETERNUS DX существуют 4 варианта реализации технологий моментальных снимков и клонов:

- *OPC* – копируются все данные с источника. Подходит для создания резервных копий с общим управлением. Обеспечивается полная резервная сохранность данных многих поколений;
- *QuickOPC* – после создания резервной копии всех данных целиком, копируются только обновленные блоки. Подходит для систем баз данных, в которых на создание резервных копий должно уходить меньше времени;
- *SnapOPC+* – копируются только внешние изменения. Помогает сократить занимаемое место на диске назначения. Подходит для резервации данных, хранимых на файловых серверах;
- *ЕС* – непрерывное копирование (зеркалирование) полностью всех данных с источника на адресат. Доступ к скопированным данным можно получить, приостановив сессию ЕС. После приостановки сессия ЕС может быть возобновлена. Копированию подлежат только те блоки, в которые внесены изменения. Это ускоряет зеркалирование.

Построение катастрофоустойчивых конфигураций

На базе дисковых массивов ETERNUS DX возможны построение катастрофоустойчивых систем и организация репликации данных. Возможна репликация синхронная и асинхронная, по интерфейсам FC или iSCSI, а также организация двунаправленной репликации и каскадных многоуровневых процессов репликации.

Дисковые системы ETERNUS DX сертифицированы на соответствие ПО Site Recovery Manager компании VMware, что позволяет автоматизировать процесс восстановления работы критических

приложений и гарантировать простоту возобновления работы систем.

Технология Thin Provisioning

Данная технология дает возможность выделения емкости логическим томам (из общего виртуального пула), превышающей объем физических дисков системы. Это позволяет существенно снизить капитальные и текущие расходы на аппаратные средства и добиться большего повышения показателя окупаемости инвестиций.

Простое и эффективное управление ИТ-инфраструктурой

Системы ETERNUS DS подходят для использования в составе широкого диапазона различных приложений в гетерогенных средах. ETERNUS SF Storage Cruiser, управляющее программное обеспечение компании Fujitsu, выполняет конфигурирование и мониторинг всех систем хранения данных на базе ETERNUS в среде SAN. Оно также поддерживает управление взаимосвязями между хранилищем данных с точки зрения хост-серверов и системой хранения данных.

Fujitsu также предлагает легко расширяемое решение для управления гетерогенными серверами, хранилищами информации и сетевыми инфраструктурами – ManageNow® Data Center Management. Оно поддерживает консолидацию различных продуктов по мониторингу и управлению. Процессы мониторинга объединяются и централизуются. Данное решение позволяет ИТ-подразделениям максимально увязать ИТ-инфраструктуру с конкретными бизнес-задачами. Все проблемы автоматически идентифицируются и затем классифицируются согласно их приоритетам, основанным на их потенциальном воздействии на ход бизнес-процессов. Поэтому коррекция ошибок с учетом конкретных правил выполняется автоматически. Четкая и понятная визуализация еще больше упрощает управление ИТ-инфраструктурой и поддерживает простоту администрирования.

ПО ManageNow® Data Center Management включает в себя серию практических усовершенствований, например управление предварительно сконфигурированными событиями и присвоение проверенных объектов бизнес-процессам и сервисам. Технология ManageNow®, благодаря высокому уровню автоматизации и консолидации управления ИТ-инфраструктурами, позволяет снизить эксплуатационные расходы до 40%.

Заключение

Наряду с законченной серией on-line дисковых систем, компания Fujitsu в рамках семейства ETERNUS выпустила на рынок линейку виртуальных ленточных библиотек и решений по архивному хранению. Весь комплекс предлагаемых продуктов позволяет сегодня реализовать решение практически любого уровня сложности – от высокопроизводительного для блочного хранения данных до легко внедряемой в ИТ-ландшафт заказчика системы резервного хранения на дисках.

Александр Яковлев,
Fujitsu Technology Solutions