

Блочная дедупликация в решениях EMC

В конце 2009 г. компания EMC расширила свою линейку решений для резервного копирования с дедупликацией данных, добавив к программному обеспечению EMC Avamar дисковые системы Data Domain. Публикация — обзор позиционирования и особенностей нового портфеля решений подразделения EMC BRS (Backup Recovery Systems).



Павел Карнаух — руководитель направления “Резервное копирование и восстановление”, EMC Россия и СНГ.



Иван Скудин — технический консультант по системам резервного копирования и восстановления, EMC Россия и СНГ.

Введение

В последнее время ИТ-службы, отвечающие за защиту данных, сталкиваются со все более жесткими требованиями к надежности хранения резервных копий и скорости восстановления данных. Традиционные решения на основе ленточных накопителей с трудом справляются с обеспечением соглашений об уровне обслуживания применительно к средам резервного копирования и восстановления.

Очевидным выходом из этой ситуации является переход к резервному копированию на диск. Опросы аналитических компаний свидетельствуют, что все большее количество пользователей начинают использо-

вать специализированные дисковые системы для повышения производительности, надежности и управляемости операций резервного копирования и восстановления.

К сожалению, нужно отметить, что в подавляющем большинстве случаев эти достаточно дорогие системы используются неэффективно. Дело в том, что “традиционное” резервное копирование подразумевает постоянное накопление избыточной информации. Это происходит из-за того, что данные в информационных системах часто дублируются, причем одни и те же блоки данных могут присутствовать в самых разных системах, например на почтовых серверах (особенно, если документы рассылаются нескольким получателям), в файловых системах различных компьютеров, в базах данных. Использование “традиционных” процессов резервного копирования увеличивает масштабы избыточности, сохраняя избыточные данные снова и снова. В результате объем хранимых резервных копий в обычной организации превышает объем продуктивных данных в 10–20 раз, при том, что различия между двумя копиями могут составлять единицы процентов.

Для решения этой проблемы уже достаточно давно с успехом используется дедупликация данных, позволяющая выделить уникальные блоки данных, сохранить их только один раз и тем самым снизить избыточность хранимых резервных копий. Такие решения присутствуют на рынке уже достаточно давно, но в последние 2–3 года их популярность стала расти лавинообразно, что позволило аналитической компании Gartner заявить, что “Развертывание систем с дедупликацией прогрессирует необычно быстро для достаточно новой технологии... Gartner рассматривает эту технологию как революционную, потому что она радикально снижает стоимость резервного копирования и восстановления с использованием дисков... игнорировать ее невозможно.”

По словам Gartner, к 2012 г. дедупликация будет использоваться в 75% всех систем резервного копирования.

Как известно, существуют две основные технологии дедупликации данных на блочном уровне — в источнике (разбиение данных на блоки и выявление уникальных бло-

ков осуществляют программные агенты, установленные на клиентах системы резервного копирования) и на целевом устройстве (устранение избыточных данных производится системой хранения резервных копий). EMC также предлагает два решения — EMC Avamar и EMC Data Domain, реализующих, соответственно, первый и второй подходы.

EMC Avamar

EMC Avamar помогает решить проблемы, связанные с традиционным резервным копированием, сделав возможным быстрое копирование и восстановление данных для локальных и удаленных систем, а также виртуализованных сред. Это решение использует патентованную технологию глобальной дедупликации в источнике для того, чтобы уменьшить объем данных до передачи в сеть. Передавая только новые, уникальные блоки данных, Avamar достигает 500-кратного снижения загрузки сети и 20-кратного снижения емкости устройств хранения данных резервного копирования.

При резервном копировании с помощью EMC Avamar патентованный алгоритм разбивает исходные данные на блоки переменной длины, пытаясь выявить как можно больше одинаковых блоков. В качестве индексов или адресов таких блоков используется 160-битный хэш-ключ, вычисляемый для каждого блока. Средний размер блока (после обязательного дополнительного сжатия) составляет 12–14 Кбайт, которые и сохраняются на сервере Avamar.

Клиенты системы Avamar передают только уникальные блоки, еще не содержащиеся в хранилище. Если сервер ранее принимал такой блок от другого клиента, то данные этого блока повторно передаваться не будут, а будет передана только метainформация о том, что такой блок встречается также и на текущем клиенте. Таким образом, реализуется глобальная дедупликация данных: в источнике (на уровне каждого клиента Avamar) и на сервере (на уровне всех клиентов системы).

Резервные копии, создаваемые Avamar, можно охарактеризовать как “виртуальные полные”, т.е. с одной стороны они

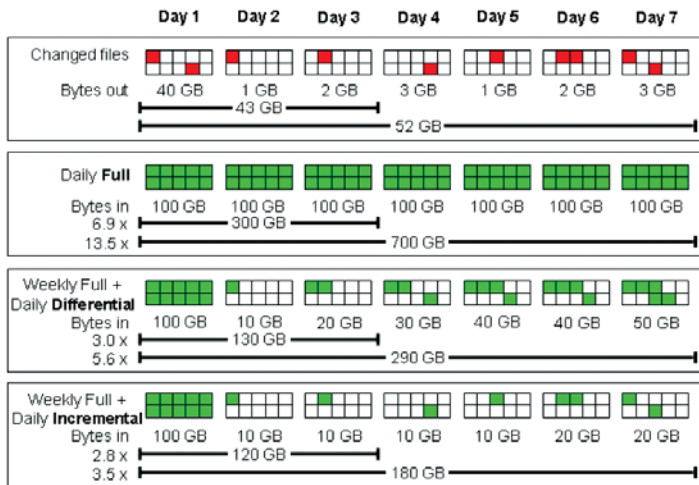


Рис. 1. В решениях с дедупликацией в источнике используются “виртуальные полные” копии, занимающие минимум места и всегда полностью готовые для восстановления в отличие от дифференциальных и инкрементальных копий.

занимают минимум дискового пространства (размер копии = размеру измененных блоков данных), а с другой — обеспечивают восстановление “за один проход” и не требуют считывания дополнительных копий, как это происходит (рис. 1) в случае дифференциального или инкрементального резервного копирования.

Обеспечение отказоустойчивости и масштабирование EMC Avamar

Серверная часть EMC Avamar строится на базе стандартных x86 серверов и архитектурно представляет собой grid-кластер. Все узлы кластера (за исключением одного — управляющего) равноправны, поэтому балансировка нагрузки для данных выполняется автоматически с привлечением всех доступных ресурсов системы Avamar, что позволяет максимально использовать все накопители, имеющиеся на серверах. Метаинформация хранится совместно с дедуплицированными данными, что позволяет проводить линейное масштабирование по емкости/производительности простым добавлением новых узлов.

Кластеры EMC Avamar используют так называемую RAIN-архитектуру (Redundant Array of Independent Nodes — избыточный массив независимых узлов). Применение RAIN в значительной степени позволило повысить надежность, масштабируемость и ремонтпригодность сервера EMC Avamar. Показатель надежности такого рода систем имеет принципиальное значение и, безусловно, должен опираться на архитектурную поддержку, а не простое дублирование компонент (что в большинстве случаев экономически не оправдано). Если произошел сбой одного из серверных узлов Avamar, можно продолжить чтение и запись в обход этого узла, что также позволяет продолжить резервное копирование и восстановление. Можно также восстановить этот узел с использованием всех серверов, входящих в систему Avamar.

Ежедневно производится внутренняя проверка контрольных точек серверов Avamar, что позволяет убедиться в том, что восстановление данных возможно во всей полноте в исходном состоянии на исходную дату. Это гарантирует высокую целостность хранимых данных.

Кроме того, для обеспечения дополнительной защиты данных от катастроф, их можно реплицировать с одного сервера Avamar на другой. При этом применяется та же технология, что и в основном решении, обеспечивающая пересылку только уникальных изменений с клиента Avamar на сервер. Репликация производится по расписанию, асинхронно, так что поврежденные данные на одной стороне не повлияет на данные, хранимые на другой стороне.

Область использования EMC Avamar

Дедупликация в источнике значительно, в десятки и даже сотни раз, снижает объем информации, передаваемой на сервер резервного копирования. Это дает возможность производить резервное копирование быстрее, тем самым снижая требования к “окну резервного копирования” в условиях ограниченных аппаратных ресурсов.

Использование EMC Avamar требует значительно меньше ресурсов (ЦПУ, памяти и каналов ввода/вывода) по сравнению с традиционным резервным копированием. Именно поэтому преимущества Avamar наиболее ярко проявляются при защите виртуализованных сред, в которых запас “свободных” ресурсов существенно меньше, чем до виртуализации, и любая дополнительная нагрузка может негативно сказаться на функционировании всех виртуальных машин.

EMC Avamar тесно интегрирован с VMware, включая:

- резервное копирование и восстановление на уровне виртуальных машин с помощью vStorage API;
- поблочное инкрементальное копирование;
- восстановление отдельных файлов из копий виртуальных машин;
- резервное копирование служебных данных vCenter.

Преимущества EMC Avamar могут быть подтверждены опытом VMware Inc., которая использует именно это решение для защиты своего ЦОД (20 Тбайт данных, более 400 виртуальных машин). После внедрения EMC Avamar среднее время копирования виртуальной машины сократилось до 8 минут, объем ежедневно перемещаемых и сохраняемых резервных копий сократился до 1% от объема продуктивных данных.

Еще одна очевидная область применения — резервное копирование данных удаленных филиалов, соединенных с центральным офисом низкоскоростными каналами передачи данных. В ряде случаев организация централизованного копирования данных филиалов возможна только с использованием EMC

Avamar; в противном случае потребуются масштабные инвестиции в наращивание мощности каналов.

В последней версии EMC Avamar 5.0 появились специализированные средства, облегчающие резервное копирование и восстановление данных с рабочих станций и ноутбуков. При этом дедуплицированные резервные копии займут в два-три раза меньше дискового пространства, чем исходные данные, а особенности лицензирования Avamar, не ограничивающие количество подключаемых клиентов, дополнительно повышают эффективность этого решения.

Стоит отметить, что кроме резервного копирования пользовательских файлов и виртуальных машин, где преимущества дедупликации в источнике абсолютно очевидны, часто имеет практический смысл применять данную технологию и для “горячего” копирования данных приложений, таких как SQL Server, Exchange, ORACLE, Lotus Domino и т.п. В состав EMC Avamar включены соответствующие модули, которые позволяют копировать данные без остановки приложений, обеспечивая и в этом случае экономию ресурсов локальной сети и емкости хранения.

Однако не все заказчики готовы перейти на новую систему резервного копирования даже при тех несомненных плюсах, которые предлагает EMC Avamar. Изменение существующих процессов и приложений является достаточно сложной задачей, поэтому растет спрос на решения для хранения резервных копий, которые могут обеспечить резкое повышение эффективности, минимально затронув существующую инфраструктуру ИТ.

В этом случае имеет смысл подробно рассмотреть технологию дедупликации на целевом устройстве, реализованную в системах EMC Data Domain.

EMC Data Domain

Дисковые системы EMC Data Domain представляют собой специализированные устройства, которые обеспечивают сохранение резервных копий, их дедупликацию, репликацию и быстрое восстановление. Они предоставляют традиционные для дисковых систем преимущества:

- высокая производительность (до 5,4 Тбайт/час суммарно для старших моделей) в сочетании с возможностью эффективной обработки низкоскоростных потоков данных, что гарантирует достижение максимально возможной скорости резервного копирования;
- кардинально более высокая по сравнению с лентами надежность, гарантированная применением стандартных отказоустойчивых решений, таких как защита дисков RAID, энерго-независимая кэш-память и т.п.;
- высокая производительность при восстановлении, обусловленная несколькими факторами: использованием высокопроизводительной дисковой подсистемы, возможностью мгновенного доступа к произвольным данным, отсутствием необходи-

мости последовательного обращения к нескольким носителям при восстановлении из инкрементальных резервных копий, а также отсутствием необходимости демультимплексирования данных.

Кроме этого, системы EMC Data Domain обладают рядом преимуществ перед обычными дисковыми системами:

- возможность создания нескольких виртуальных библиотек и виртуальных накопителей позволяет спроектировать архитектуру системы резервного копирования оптимальным образом (в случае необходимости разделяя дисковую систему между различными приложениями), а возможность монтирования носителей в любом накопителе существенно облегчает задачу клонирования информации на ленту для долговременного хранения;
- возможность эмуляции как ленточных библиотек, так и файлового сервера. Во втором случае клиенты системы резервного копирования, не подключенные к сети SAN, могут сохранять резервные копии своих данных по локальной сети (протоколы CIFS, NFS и OST).

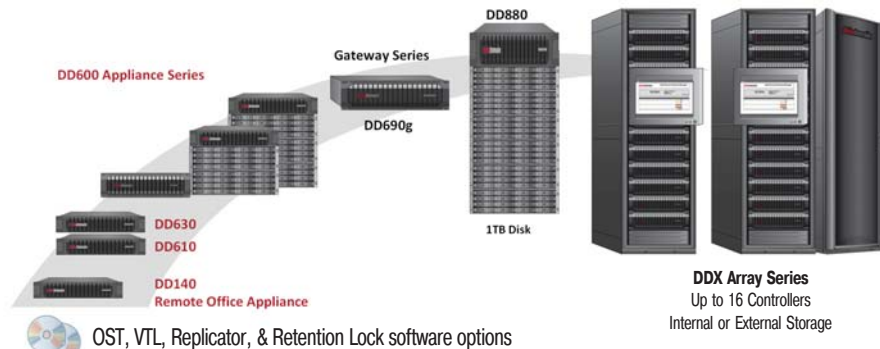
Поддержка дедупликации в дисковых системах EMC Data Domain

Важная особенность, которая отличает продукты Data Domain от прочих систем — это то, что поддержка дедупликации закладывалась в них изначально, и дополнительный функционал — репликация, эмуляция ленточных библиотек (VTL) и т.п. — разрабатывался с учетом этой технологии. В этих системах используется наиболее эффективный алгоритм дедупликации на уровне блоков данных переменной длины, позволяющий сократить объем хранилища резервных копий в десятки раз по сравнению с традиционными решениями на лентах или дисках.

EMC Data Domain поддерживает дедупликацию “на лету” (inline). На практике это означает, что системы Data Domain используют дисковую емкость в разы более эффективно, чем системы, использующие отложенную дедупликацию. Системы Data Domain эффективнее реплицируют данные, так как процесс репликации может начаться и завершиться практически одновременно с процессом резервного копирования. Системы Data Domain проще в управлении и настройке, так как пользователям не нужно заниматься сложным планированием расписания различных задач, следить за тем, чтобы они не “отъедали” друг у друга ресурсы и т.п.

В конечном счете, системы EMC Data Domain обеспечивают предсказуемую скорость резервного копирования и восстановления, в отличие от систем с отложенной дедупликацией, производительность которых невозможно предсказать в случае, если дедупликация не успела закончиться до начала резервного копирования или восстановления.

Процесс дедупликации является полностью прозрачным для серверов приложений и ПО резервного копирования.



Характеристика	DD140	DD610	DD630	DD565	DD660	DD690	DD880	DDX Array
Скорость (GB/hr)	450	675	1.1 TB/hr	1.1 TB/hr	2 TB/hr	2.7 TB/hr	5.4 TB/hr	86.4 TB/hr
Логический объем (TB)	17-43	75-195	165-420	320-810	520-1.31 PB	710-1.7 PB	1.4-3.5 PB	22.6-56.7 PB
Полезный объем (TB)	.86	Up to 3.98	Up to 8.4	Up to 16.2	Up to 26.1	Up to 35.3	Up to 71	Up to 1.13 PB

Рис. 2. Линейка решений EMC Data Domain.

Особенности дисковых систем EMC Data Domain

В отличие от большинства производителей, вынужденных устанавливать большое количество дисков для достижения приемлемой производительности, системы Data Domain были изначально ориентированы на максимальное использование возможностей процессоров и оперативной памяти. Технология SISL (Stream Informed Segment Layout) оптимизирует процесс дедупликации таким образом, что выявление дубликатов практически не требует обращения к диску, и операции проводятся в оперативной памяти. Таким образом, скорость дедупликации зависит не от количества дисков, а от производительности процессоров, что обеспечивает максимальную производительность даже при использовании SATA-дисков.

Для систем с поддержкой дедупликации крайне важно обеспечить надежное хранение данных. В EMC Data Domain реализована архитектура DIA (Data Involuntarily Architecture), обеспечивающая сквозную проверку целостности данных, обнаружение сбоев и восстановление работоспособности систем.

Поддержка репликации

С 2004 г. системы Data Domain обеспечивают построение катастрофоустойчивых решений с помощью репликации данных на резервную площадку по сетям IP. При репликации передаются только измененные блоки данных, что помогает минимизировать требования каналу передачи данных. Поддерживаются различные схемы, в т.ч. “один к одному”, “многие к одному”, двунаправленная репликация и т.п.

Использование схемы репликации “многие к одному” позволяет сформировать гибкое и масштабируемое решение для крупных организаций с географически распределенными филиалами. Теперь появилась возможность за счет использования глобальных сетей с низкой пропускной способностью упростить и удешевить инфраструктуру аварийного восстановления данных и одновремен-

но с этим повысить степень готовности организации к восстановлению данных в случае сбоев.

В сентябре 2009 г. решения EMC Data Domain были расширены опцией каскадной репликации. Эта опция позволяет оптимизировать стратегию аварийного восстановления, используя высокоэффективную репликацию для передачи дедуплицированных резервных копий на несколько резервных площадок.

Область использования систем EMC Data Domain

Системы EMC Data Domain обеспечивают одни из лучших в отрасли показатели производительности при резервном копировании и восстановлении: как для одного потока данных (особенно важно для восстановления), так и суммарно по системе в целом. При этом обеспечивается полная независимость дедупликации от типов данных и приложений резервного копирования.

EMC предлагает широкую продуктовую линейку систем Data Domain, в которую входят 7 моделей — от DD140, ориентированных на использование в удаленных офисах, до DD880, предназначенных для использования в крупных ЦОД (рис. 2). Большинство заказчиков используют системы Data Domain для резервного копирования всех типов данных, включая многотерабайтные БД.

Использование технологий дедупликации позволяет перейти к “безленточному” резервному копированию данных, существенно повысить надежность хранения резервных копий и минимизировать затраты на внедрение репликации резервных копий в удаленный ЦОД.

Табл. 1. Сравнение решений EMC Avamar и EMC Data Domain.

Характеристика	Система резервного копирования EMC Avamar	Система хранения резервных копий EMC Data Domain
	В источнике	На целевом устройстве
Способ дедупликации	Да	Да
Сокращение объема хранения	Да	Нет
Сокращение сетевого трафика	Да	Нет
Сокращение нагрузки на серверы	Да	Да
Повышение надежности хранения	Да	Да
Сокращение времени восстановления	Да	Да
Интерфейсы/протоколы	GigE	GigE/10GigE (NFS, CIFS, OST), FC (VTL)
Скорость восстановления данных	Десятки Мбайт/с	Сотни Мбайт/с
Интеграция с “традиционными” системами резервного копирования	EMC NetWorker	EMC NetWorker, Symantec NetBackup, HP Data Protector, IBM TSM и т.п.
Рекомендуемая область применения	Резервное копирование виртуальных сред, удаленных филиалов, файловых серверов, рабочих станций и ноутбуков	Резервное копирование любых ресурсов, включая многотерабайтные БД

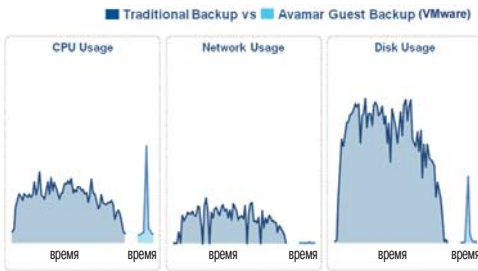


Рис. 3. Использование ресурсов клиентом EMC Avamar, например, выполняющегося как guest backup в среде VMware, с традиционным бэкапом больше, но много короче.

Особенности позиционирования EMC Avamar и EMC Data Domain

И системы EMC Avamar и системы EMC Data Domain поддерживают дедупликацию резервных копий на уровне блоков переменной длины. Различия между ними определяются прежде всего способом дедупликации — в источнике или на целевом устройстве, и возможностью интеграции в существующую систему резервного копирования. В табл. 2 приведены сравнительные характеристики этих решений.

В системах EMC Avamar восстановление данных производится только по IP-сети (Gigabit Ethernet), в системах EMC Data Domain — как по IP-сети (Gigabit Ethernet, 10Gigabit Ethernet), так и по сети хранения FC SAN. С этой точки зрения рекомендуется использовать EMC Avamar для резервного копирования относительно небольших серверов

приложений и СУБД — не более 1 Тбайт данных на одном клиенте системы Avamar.

При проектировании систем резервного копирования файловых серверов целесообразно учитывать не скорость резервного копирования (так как объем передаваемых данных составит менее 1% от общего объема), а количество файлов. Средняя скорость сканирования файловой системы при копировании с помощью Avamar составляет 1 млн файлов в час. Учитывая, что большинство серверов содержат значительно меньшее количество файлов, резервное копирование с помощью Avamar, как правило, проходит в несколько раз быстрее, чем при использовании “традиционных” решений. Это касается и нагрузки на ресурсы клиентов системы резервного копирования (процессор, диск, сеть); например, нагрузка на процессор будет несколько выше, чем обычно, но период нагрузки будет намного короче (рис. 3).

Дисковые системы EMC Data Domain применяются совместно с “традиционным” ПО резервного копирования. Они легко интегрируются как с EMC NetWorker, так и с решениями других производителей. В отличие от Avamar эти системы не обеспечивают сокращения нагрузки на сеть и ресурсы клиентов, но обладают рядом других преимуществ — в частности, повышенной производительностью при восстановлении по сети FC SAN или 10GigE, что позволяет позиционировать их для хранения резервных копий многотерабайтных БД.

Заключение

Согласно отчету Gartner¹⁾ в настоящее время в портфеле EMC две наиболее зрелые технологии дедупликации из предлагаемых на рынке: Avamar с дедупликацией на стороне клиента и Data Domain с дедупликацией на целевом устройстве.

Avamar был анонсирован в 2002 г.; в настоящее время поставляется 5-я версия этого решения. Его также отличает эффективность, надежность и стабильность работы. EMC Avamar установлен более чем у 3000 заказчиков, что примерно на порядок превышает показатели ближайшего конкурента.

Системы Data Domain поставляются на рынок с 2003 г.; их отличает отлаженность технологий и самая большая доля рынка (54% по оценке Taneja Group²⁾). В мире установлено более 10 000 систем Data Domain, из них 2500 — в Европе.

В соответствии с другим отчетом (опубликованном в ноябре 2009 г.) — рейтингу Technology Fast 500, Deloitte LLP, показывающему 500 самых быстрорастущих компаний Северной Америки с ежегодно удваивающимся оборотом в период с 2004 г. по 2008 г., компания Data Domain заняла в нем 7 место. Это свидетельствует о очень большом интересе бизнеса и госкомпаний к технологиям дедупликации и конкретно к решениям Data Domain, а также о высоком потенциале этих решений и о росте спроса на технологии дедупликации в ближайшей перспективе.

Павел Карнаух, Иван Скудин, компания EMC Россия и СНГ.

Прогнозы футуролога Дэйва Эванса

Декабрь 2009 г. — В связи с 25-летием Cisco главный футуролог консалтингового подразделения Cisco IBSG Дэйв Эванс (Dave Evans) обнародовал свой прогноз дальнейшего развития технологий. По его мнению:

- К 2029 году за \$100 долл. можно будет купить СХД емкостью 11 Пбайт. Такого объема электронной памяти будет достаточно, чтобы круглосуточно проигрывать видео DVD-качества в течение 600 лет.
- В течение двух следующих лет объем информации во Всемирной сети будет удваиваться каждые 11 часов.
- К 2020 году каждый житель нашей планеты будет в среднем хранить 130 Тбайт персональных данных (сегодня этот объем равен 128 Гбайт).



- Через 5 лет любая поверхность сможет выполнять функции дисплея.
- В ближайшие 2 года объем информации в нашем мире будет ежегодно увеличиваться в шесть раз, а объем корпоративных данных в тот же период будет ежегодно возрастать в 50 раз.
- К началу 2010 года к Сети окажутся подключены 35 млрд различных устройств, т.е. почти по 6 устройств на каждого жителя нашей планеты.
- В предстоящие 10 лет скорость передачи данных в домашних сетях увеличится в 20 раз.
- К 2013 году ежемесячный объем трафика в беспроводных сетях составит 400 петабайт (сегодня весь мировой сетевой трафик составляет 9 экзбайт в месяц).
- К концу 2010 года на каждого жителя нашей планеты будет приходиться по 1 млрд транзисторов стоимостью одна миллионная американского цента каждый.
- Первый коммерческий квантовый компьютер появится к середине 2020 года.
- К 2015 году человечество будет ежегодно создавать контент, объем которого в 92,5 млн раз превышает объем информации, хранящейся в библиотеке Конгресса США.
- К 2015 году повсеместно распространится видеосвязь и она будет генерировать 400 экзбайт трафика, что в 20 млн раз превышает объем информации, хранящейся в библиотеке Конгресса США.

- К 2015 году объем данных, которые будут генерировать телефонная связь, Интернет, электронная почта, фото- и музыкальные файлы, составит 50 экзбайт.
- С внедрением протокола IPv6 в Интернете появится такое количество электронных адресов, что каждую из известных человечеству звезд во Вселенной можно будет снабдить 4,8 трлн адресов.

8Gb FC HBA-решение для шифрования данных “на лету”

Октябрь 2009 г. — Компания Emulex объявила, что будет совместно с IBM поставлять решение для шифрования данных на базе Emulex Secure 8Gb Encryption HBA технологии (<http://www.emulex.com/products/strategic-direction/emulexsecure-8gb-encryption-hba.html>) и IBM Tivoli Key Lifecycle Manager.

Это решение на основе 256-битной AES технологии позволит обеспечить безопасность данных (шифруя/дешифруя их “на лету”) как при передаче данных, так и при хранении, сокращая капитальные затраты до 50–90% в сравнении как с программным шифрованием, так и аппаратным на уровне SAN.

1) “EMC Acquires Data Domain, Becomes Deduplication Leader, Signals Deduplication as a “Must Have” Capability”, Dave Russell (Gartner, 25 August 2009/ID Number: G00170233).
2) Taneja Group, Sept. 2008