

# Можно ли “забыть”, где хранятся данные?

*В конце I кв. с.г. компания EMC анонсировала третье поколение (ver. 2.2) своего продукта – Centera, ориентированной на хранение т.н. неструктурированного контента. С момента появления первой версии (около двух лет назад) Centera претерпела значительные расширения, и на настоящий момент – это практически единственная система хранения в своей нише, а также – единственное решение, сертифицированное Комиссией по ценным бумагам США и Министерством обороны США, для хранения данных, искажение которых не допускается.*

## Введение

По данным аналитических компаний (Gartner, IDC), доля неструктурированной информации уже сейчас составляет от 70-80% общего объема при стабильном росте в ближайшей перспективе. И хотя отношение к данному типу информации не всегда однозначно, показательны в этом отношении позиции правительства США и многих европейских государств, которые на законодательном уровне обязывают компании хранить в электронном виде корпоративную информацию определенного вида в течение длительного времени. При этом регламентируются не только сама информация, но и способ ее хранения и тип носителей. В частности например, обязательным является не только наличие резервного центра, но и его удаленность от основного – не ближе 100 км. Эти и другие требования к хранению неструктурированного (или фиксированного) контента, как и сам контент, накладывают дополнительные условия, предъявляемые к системам хранения и к организации процесса хранения в целом.

Под фиксированным контентом (Fixed Content, или “ссылочная информация” – Reference Information) понимаются цифровые данные, которые были созданы в результате какой-либо деятельности, но теперь являются статическими (т.е. никогда не будут изменяться). Фиксированный контент хранится в течение длительного периода (многих лет или десятилетий), который зависит от типа требований к данным документам, установленных правительством, рамками исследований или политикой компании. В качестве примеров можно при-

вести цифровые документы (финансовые отчеты, контракты, платежные документы), фотографии, медицинские снимки (рентгеновские, MRI и др.), медианосители (включая видеоновости, записи видеонаблюдений), аудио (включая речевую почту или MP3-файлы), геофизические данные (включая сейсмические, астрономические или географические), другие электронные активы (типа книг, руководств, писем и вложений, презентаций или электронных таблиц) и CAD/CAM-проекты. Все они имеют общность в том, что после того как контент получен (например, видео произведено, отредактировано и подготовлено для распространения), его хранят очень длительное время. К нему могут обращаться тысячи раз тысячи пользователей, но его содержимое остается неизменным.

Большая часть информации, хранящейся сегодня в виде фиксированного контента, доступна в режиме off-line или near-line. Запросы, dictуемые временем, выдвигают требования по доступу к нему в режиме on-line – предоставление нового уровня сервиса пользователям. Например, большинство клиентов готовы платить за то, чтобы иметь возможность онлайнового доступа к их платежным документам.

Резюмируя особенности фиксированного контента в сравнении с изменяемыми данными, можно отметить следующие особенности последних:

- непрерывное и интенсивное изменение данных в результате работы транзакционных приложений;
- ориентация, в основном, на блоковый доступ к данным;

- большое количество “мелких” операций чтения/записи;
- типичные примеры приложений: СУБД, OLAP, OLTP, ERP-системы, учетно-финансовые системы;
- наилучшая реализация на системах хранения типа: SAN, DAS, NAS;
- жизненный цикл информации – часы, дни, месяцы.

Для фиксированного контента характерным является:

- неизменность данных после их создания и невысокая активность их использования;
- единица информации – объект/файл;
- большое количество операций по чтению/записи двоичных объектов;
- независимость от файловой системы, наибольший уровень масштабируемости и надежности, возможность одновременного прямого доступа к информации большого числа пользователей;
- жизненный цикл информации – годы и десятилетия.

Подводя итог рассмотрению особенностей хранения фиксированного контента, можно выделить следующие обобщенные требования, которым должна отвечать ориентированная для этого современная система хранения. При очень длительном сроке хранения данных система должна отвечать повышенным требованиям надежности; иметь минимальные затраты на ТСО, т.е. быть в максимальной степени самовосстановливаемой, самоуправляемой; автоматически поддерживать удаленный резервный

центр; обеспечивать совместную работу нескольких поколений аппаратных средств; максимально масштабироваться по объему и пропускной способности при минимальных затратах; осуществлять автоматическую реорганизацию данных при необходимости; быть в максимальной степени открытой платформой для совместного одновременного использования многими пользователями.

## Система хранения EMC Centera

Появившаяся на рынке около двух лет назад (*SN № 5/14, 2002*) система хранения EMC Centera остается по настоящее время уникальной разработкой, оптимизированной для хранения больших массивов критически важных данных в течение многих лет с минимальными затратами на управление и модернизацию.

Основные проблемы, возникающие при хранении больших объемов информации в течение долгого времени, связаны с необходимостью постоянно решать целый ряд задач по расширению емкости, увеличению производительности, устранению неисправностей, миграции с устаревших хранилищ на новые. EMC Centera позволяет решить все эти проблемы благодаря принципиально новому подходу к хранению данных.

Изначально построенная для хранения миллионов объектов, система Centera представляет собой симбиоз объектной базы данных CentraStar и аппаратного комплекса по хранению информации. Сконцентрировав в себе, как емкость, так и интеллектуальность по автоматизированному управлению данными, Centera самостоятельно решает большую часть рутинных задач, обычно связанных с хранением больших объемов информации.

При работе с Centera приложения используют интерфейс программирования Centera API, позволяющий сохранять, получать и удалять объекты из хранилища. При этом задачи по оптимальному размещению данных, поиску объектов, устраниению неисправностей, балансировке нагрузки, расширению емкости и т.д. решаются системой Centera самостоятельно.

## Архитектура системы хранения EMC Centera

Аппаратная составляющая системы хранения Centera представляет собой совокупность узлов (рис. 1), отвечающих за обработку запросов приложений (Access Node – AN)

и за хранение данных (Storage Node – SN), объединенных в RAIN-архитектуру (Redundant Array of Independent Nodes). В системе Centera используется не менее двух узлов для обработки запросов, причем количество таких узлов обычно увеличивается с ростом емкости Centera, что позволяет увеличивать производительность пропорционально росту емкости. Каждый узел хранения информации содержит 1,2 Тбайт незащищенного дискового пространства и соединен со всеми остальными узлами Centera через дублированную внутреннюю сеть Gigabit Ethernet, построенную на Layer 3 коммутаторах.

Минимальное количество узлов в Centera составляет 8. В одном кабинете Centera может быть установлено до 32 узлов с возможностью масштабирования до 8 кабинетов в составе одной системы хранения. В дальнейшем ожидается поддержка большего количества кабинетов, причем любых из когда-либо выпущенных поколений Centera, в одной системе хранения.

## Обеспечение сохранности данных

В связи с тем, что Centera решает сразу обе задачи, как хранения, так и управления данными, на принципиально другом уровне обеспечивается защита данных как от аппаратных сбоев, так и от ошибок администратора или приложения, использующего систему Centera.

Каждый сохраняемый на Centera объект обязательно дублируется на двух различных узлах при использовании Content Protection – Mirroring (CPM). Причем выбираются те узлы, которые наименее загружены и имеют больше свободного пространства. При защите Content Protection – Parity (CPP) объект разделяется на шесть частей и записывается на шесть различных узлов с записью контрольной суммы (по аналогии с RAID 5) на седьмой узел, что так же, как и CPM, позволяет полностью защититься от сбоя любого узла.

Полезная защищенная емкость одного кабинета Centera составляет от 3,58 Тбайт до 19,1 Тбайт в случае CPM защиты и от 11,92 Тбайт до 31,79 Тбайт в случае CPP защиты. Восемь кабинетов позволяют иметь полезную защищенную емкость в 254 Тбайт в одной системе хранения. Интерфейс программирования Centera API одновременно адресует несколько систем хранения Centera с испытанный на практике общей емкостью доступного пространства в 1 Пбайт.

В случае выхода из строя узла, остальные узлы, содержащие объекты, которые также хранились на сломавшемся узле, автоматически сделают еще одну копию объектов. Таким образом, в процессе копирования 1,2 Тбайт данных участвует максимальное количество узлов, что позволяет восстановить необходимую избыточность данных за минимальное время без заметного влияния на производительность хранилища.

При сохранении объекта приложение может указать минимальный период хранения (retention period). В этом случае объект не может быть удален приложением до окончания данного периода, что позволяет гарантировать сохранность информации в течение заданного времени даже при ошибках оператора, сбоях приложения или диверсии.

Дополнительную защиту обеспечивает проверка целостности данных, которая выполняется регулярно для всех хранимых объектов. При такой проверке сверяется содержимое объекта с контрольной суммой и с уникальным идентификатором. При искашении объекта на одном из узлов, он автоматически восстанавливается при помощи CPM или CPP защиты.

Еще один уровень защиты обеспечивает возможность репликации данных между несколькими системами Centera, которые могут быть расположены в различных центрах обработки. Такая репликация, обеспечиваемая CentraStar, позволяет защитить данные от масштабных аварий и катастроф. При этом максимально упрощается процесс получения данных в резервном центре, так как они будут автоматически запрошены через интерфейс Centera API с резервной системы при недоступности основной системы Centera.

## Защита инвестиций и расширение емкости

В составе одной системы хранения Centera могут одновременно использоваться все, когда-либо выпущенные, поколения узлов Centera. При этом CentraStar будет учитывать разницу в производительности и емкости различных узлов при размещении данных. Приложения, использующие очень удобный и простой интерфейс программирования, защищены от технологических изменений благодаря тому, что работа системы Centera абсолютно не зависит от того, какие узлы в ней используются. Более того, даже в случае появления в будущем принципиально других узлов, использующих, например, твердотельную память, а не дисковые накопители, не потребуется вносить ни одного изменения в приложения, а эти новые узлы можно будет использовать совместно с уже имеющимися.

При добавлении узлов в кабинет или добавлении нового кабинета Centera автоматически определяет новое доступное пространство и начинает немедленно его использовать. Таким образом, полностью отсутствуют проблемы, связанные с расширением логических томов в системах хранения SAN или файловых систем в NAS-хранилищах.

## Адресация данных

Система хранения EMC Centera является системой с адресацией данных по их содержимому, или CAS-системой (Content Addressed Storage).

При сохранении объекта система Centera вычисляет уникальный идентификатор объекта по алгоритму MD5 компании RSA Data Security с использованием содержимого объекта. При сохранении нескольких объектов с идентичным содержимым, Centera будет содержать только одну копию объекта, что позволит сократить объем используемого пространства.

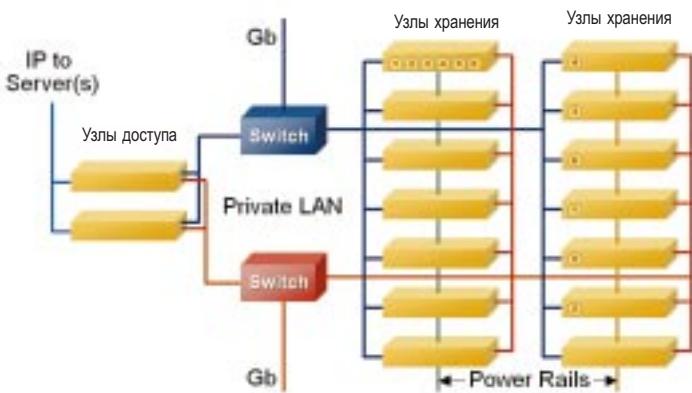


Рис. 1. Архитектурно Centera состоит из узлов доступа (Access Node) и узлов хранения (Storage Node), соединяемых между собой через два коммутатора.

Для получения объекта или его удаления приложение использует уникальный идентификатор объекта. Очевидно, что при попытке злоумышленника незаметно исказить данные на системе хранения, приложение по-прежнему будет получать по своему идентификатору неискаженные данные, так как искаженному объекту будет соответствовать совершенно другой незнакомый приложению идентификатор. Благодаря такой защите данных система Centera является на текущий момент единственной системой хранения, сертифицированной Комиссией по ценным бумагам США и Министерством обороны США для хранения данных, искажение которых не допускается (финансовая и секретная информация).

При сохранении объекта приложение также может сохранить в Centera дополнительные свойства в формате XML, по которым в дальнейшем может производиться поиск и фильтрация выдаваемых объектов, что особенно важно при восстановлении идентификаторов объектов, если они были по какой-то причине потеряны приложением.

Каждый идентификатор занимает 128 бит. Если средний размер объекта составляет, например, 1 Мбайт, то при хранении 1 Гбайт информации на Centera общий объем идентификаторов составит всего лишь 16 Гбайт. Только для этих данных нужно производить регулярное резервное копирование, что при таких маленьких объемах не составляет труда. Резервное копирование данных, хранимых на Centera, также может производиться при помощи протокола NDMP, но практической необходимости в этом нет, особенно при использовании минимального срока хранения объектов (retention period).

### Производительность EMC Centera

Из-за возможности свободно определять количество узлов обработки запросов приложений, производительность Centera может расти линейно при росте объема. При этом любой из хранимых объектов будет выдаваться приложению незамедлительно благодаря использованию дисковых накопителей для хранения данных и, как следствие, отсутствию в системе каких-либо роботов, перевставляющих ленты или диски.

Все потоки в Centera распределяются равномерно по всем доступным узлам, но для сбалансированности нагрузки между узлами доступа и хранения важно выдерживать правильное соотношение между ними:

- 2 узла доступа на 8 узлов Centera (6 SN, 2 AN);
- 4 узла доступа на 16 или 32 узла Centera;
- от 2 до 4 узлов доступа на каждый дополнительный кабинет в кластере;
- в целом – 1 узел доступа на 7 узлов хранения.

Добавление узлов доступа, выше рекомендованных, улучшает производительность Centera по чтению от 20% до 70% в зависимости от размера файла (объекта), и чем меньше размер файла, тем меньше прирост производительности (рис. 2). Это объясняется тем, что вследствие одинакового объема

Оптимальная интенсивность потока запросов к 32-узловой Centera (28 node storage, 4 node access) по чтению/записи при 80 параллельных процессах

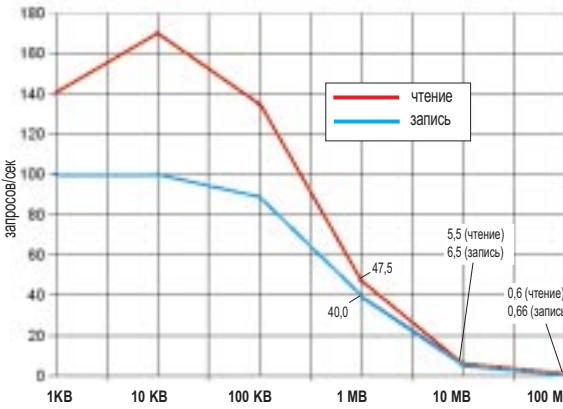


Рис. 2. Изменение пропускной способности Centera (запросов/секунду) в зависимости от размера файла и типа операции.

накладных затрат независимо от размера передаваемого объекта (файла), общая пропускная способность снижается при уменьшении размера объекта. При этом следует учитывать то, что на каждый объект записи в Centera пишется четыре объекта: (сам файл + его описание) x 2 (при зеркалировании).

Важной особенностью архитектуры Centera является то, что она поддерживает множество параллельных потоков, поэтому для увеличения общей пропускной способности важно использовать многопоточные приложения. Однако важно оптимизировать параллельные процессы в системе, чтобы существенным образом не росли задержки из-за конкуренции за ресурсы (рис. 3).

Все узлы доступа Centera могут быть подключены к коммутатору и работают параллельно.

Производительность Centera часто оказывается намного выше, чем интенсивность запросов объектов сервера приложений, которая обычно ограничена производительностью конкретного сервера или количеством обращений пользователей в единицу времени. Например, система Centera с 28 узлами для хранения данных и 4 узлами обработки запросов позволяет получать более полутора миллиона документов объемом 100 Кбайт/час, что обычно превосходит потребности даже крупных организаций.

Следует учитывать и то, что при чтении сервер приложения посылает приложению только CAS-адрес, дальнейшее считывание объекта происходит уже без участия сервера, что в ряде случаев позволяет во многом снизить требования к серверу.

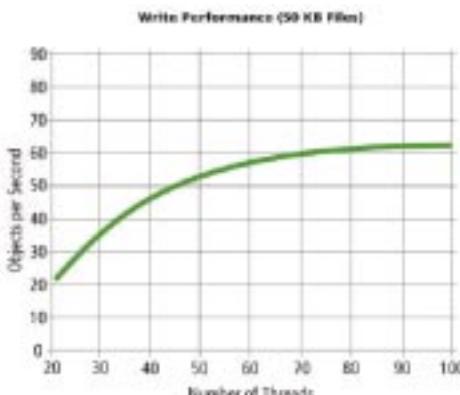


Рис. 3. Оптимальное соотношение между числом параллельных потоков и интенсивностью обращений на запись.

### Управление EMC Centera

Система Centera самостоятельно осуществляет конфигурирование своих узлов, балансирует нагрузку между ними, при необходимости мигрируя данные между узлами хранения, автоматически восстанавливает необходимую избыточность при любых сбоях. Управление данными производится самими приложениями, которые в соответствии со своей логикой создают и удаляют объекты.

Таким образом, администрирование системы Centera сводится лишь к мониторингу доступного свободного пространства и наблюдению за статистикой по аппаратным сбоям и производительности.

Графический интерфейс мониторинга Centera, называемый Centera Viewer, позволяет осуществлять следующие действия с рабочей станции под управлением Windows или Solaris:

- наблюдение за нагрузкой системы хранения в общем или конкретных узлов: количество записываемых и считываемых объектов, размеры объектов, количество мегабайт в секунду на чтение и запись;
- наблюдение за объемом реплицируемых данных между различными системами Centera;
- наблюдение за величиной используемой и свободной емкости;
- проверка логов событий по аппаратным сбоям.

Кроме этого, система хранения Centera может предупреждать оператора об аппаратных сбоях или нехватке емкости посредством использования протокола SNMP.

### Интеграция с приложениями

На данный момент более 100 различных приложений поддерживают Centera API и позволяют хранить данные на Centera. Спектр приложений очень широк: от специализированных приложений по документообороту, хранению рентгеновских снимков, данных геологоразведки – до приложений для архивации данных SAP, Exchange, Lotus Notes и построения иерархических систем хранения данных. Несколько программных платформ, выпускаемых EMC, также поддерживают EMC Centera: HSM-системы EMC Avalon и Legato DiskXtender, а также ECM-система Documentum. При этом “объект” системы Centera для каждого из приложений представляет собой именно те объекты, которыми оперирует приложение: фотография, документ, файл и т.д.

При использовании Documentum полностью реализуется функциональность Centera, включая возможность задания минимального периода хранения (retention period) для хранимых документов. Настройка интеграции осуществляется прямо из графического интерфейса Documentum при его первичной установке или при дальнейшем изменении конфигурации, что позволяет администратору Documentum использовать Centera без необходимости в дополнительном обучении. Для определения того, какие документы будут храниться на Centera, используются стандартные правила Documentum, что позволяет осуществлять гибкую настройку.

Табл. 1. Сравнение особенностей Centera с двумя альтернативными решениями

	EMC Centera	Ленточная библиотека	Оптическая библиотека
<b>Способ доступа к данным</b>	Режим онлайн. Все данные на дисковых накопителях	Полуонлайн (если лента уже загружена) или оффлайн (если требуется загрузка ленты роботом)	Полуонлайн (если диск уже загружен) или оффлайн (если требуется загрузка диска роботом)
<b>Управление файловыми системами и томами</b>	Не требуется. Автоматическое управление.	Требуется	Требуется
<b>Автоматическое использование добавленной емкости</b>	Да	Нет. Требуются работы по распределению данных между несколькими библиотеками.	Нет. Требуются работы по распределению данных между несколькими библиотеками.
<b>Автоматическая балансировка нагрузки</b>	Да	Нет	Нет
<b>Задачи по регулярному обслуживанию и проверке носителей информации</b>	Не требуется.	Требуется. Вручную.	Требуется. Вручную.
<b>Регулярная проверка целостности данных</b>	Да	Нет	Нет
<b>Независимость данных от местоположения</b>	Да	Нет. Конкретная лента.	Нет. Конкретный диск.
<b>Автоматизированная и незаметная для приложений миграция данных на новые устройства</b>	Да	Нет. Требуется вручную переносить данные с устаревшей библиотеки на новую (если, конечно, в старой библиотеке еще читаются ленты).	Нет. Требуется вручную переносить данные с устаревшей библиотеки на новую (если, конечно, в старой библиотеке еще читаются диски).
<b>Аутентификация</b>	Для каждого объекта по его адресу	Нет	Нет
<b>Гарантия от преждевременного уничтожения</b>	Да, для каждого объекта индивидуально	На уровне носителя	На уровне носителя
<b>Удаление содержимого</b>	Каждый объект удаляется индивидуально	На уровне носителя	На уровне носителя
<b>Единая точка отказа</b>	Отсутствует. Полное дублирование	Робот	Робот
<b>Автоматическое восстановление после сбоя</b>	Да	Нет	Нет
<b>Защита от катастроф</b>	Репликация на другую систему EMC Centera	Нет	Нет

В случае использования приложения, которое не поддерживает EMC Centera API, существует несколько простых способов интеграции. Например, при помощи HSM-систем, когда приложение работает с огромными файловыми системами, большая часть которых на самом деле хранится на Centera. Или при помощи NAS-шлюза Centera Application Gateway, позволяющего осуществлять доступ к данным по протоколам CIFS и NFS.

Кроме того, интерфейс программирования Centera API может быть предоставлен заказчику для использования в самостоятельно разработанных приложениях. Centera API поддерживает C/C++ и Java для операционных систем Sun Solaris, HP HP-UX, IBM AIX, Linux, SGI IRIX, Windows NT/2000/XP. А начиная с февраля 2004 года также поддерживается IBM z/OS.

Одна система Centera может одновременно использоваться различными приложениями. Для каждого из приложений можно настро-

ить соответствующий уровень доступа, что не позволит получить неавторизованный доступ к данным.

Сравнительная оценка особенностей EMC Centera перед такими альтернативными методами хранения больших объемов информации, как использование оптических или ленточных библиотек, приведены в табл. 1.

### TCO

Общая стоимость обслуживания или эксплуатационных затрат (Total Cost of Ownership – TCO) – один из основных факторов, определяющих выбор системы для долговременного хранения. Так, для Centera, по стандартным нормативным требованиям, при отсутствии ее переполнения появление инженерного персонала для профилактических мероприятий требуется один раз в полгода. При этом требуется гораздо более меньшая реактивность персонала при возникновении неисправностей, поскольку даже многочисленные отказы приводят

только к некоторой деградации системы, а не к ее отказу в целом.

За счет возможности одновременного использования нескольких поколений узлов отсутствует необходимость наличия склада запасных частей или необходимость приобретения у поставщика комплектующих, уже снятых с производства (которые, как правило, по своим характеристикам не только намного уступают последним релизам, но даже стоят больше).

### Заключение

*Любое общество свое будущее строит на своей истории в прошлом, и это неизбежная закономерность. В большинстве случаев это обусловлено не только требованиями бизнеса и/или компаний, но и общества в целом. И в данном контексте правильный выбор технологической основы для решения проблемы хранения фиксированной информации во многом определяет стабильность развития государства.*

## Хранение тоже может быть вечным

EMC Centera 3-го поколения

- Поддержка всех поколений компонент в одной системе
- Полная независимость приложений от физической реализации компонент системы хранения
- Полная неизменность хранимых данных
- Автоматическое внутреннее и удаленное реплицирование данных
- Максимальное масштабирование системы по емкости и производительности
- Максимальное самоконфигурирование и самовосстановление
- Профилактическое обслуживание не чаще одного раза в полгода

**EMC<sup>2</sup>**  
where information lives

Компания CLASSICA CIS – комплексное проектирование и создание систем высокой доступности и отказоустойчивых комплексов на базе кластерных технологий

Москва, ул. Гончарная 12, стр. 1  
тел.: (095) 796-9040, факс: (095) 915-8543  
e-mail: storage@classics.ru, www.classics.ru

Официальный партнер  
EMC Corporation

**КЛАССИКА**  
Computer Information Solutions