ПЕРЕДМОВА

Випуск дайджесту присвячено досвіду установ світу щодо зберігання і використання електронної інформації та мікрофільмів в сучасному інформаційному суспільстві, наведено технічні характеристики систем зберігання електронної інформації.

У публікації «Микрография и архивное хранение документов» розповідається що світ знов звернувся до випробуваної та надійної технології довготривалого зберігання інформації – мікрографії.

У публікації «Фонд изданий на микроформах» розповідається про склад та порядок користування фондом видань на мікроформах Російської національної бібліотеки.

У публікації «FAQ по восстановлению повреждённых изображений» розповідається як відновити втрачене зображення з дисків та флеш карт.

У публікації «Активные архивы в хранении данных» розповідається про визначені в окремий клас систем зберігання активні архіви, які забезпечують довготривале зберігання інформації з наданням активного доступу до любої частини архіву, в режимі реального часу.

У публікації «Опыт оцифровки архивных документов в Центральном государственном архиве города Москвы» розповідається про практику оцифровування архівних документів у центральних архівах міста Москви та методологічні проблеми довготривалого зберігання електронних документів.

У публікації «Как построить эффективный электронный архив на вашем предприятии» розглянуто декілька най розповсюджених технологій створення та ведення електронного архіву і надано висновок, яка з них приносить найбільший ефект.

У публікації «Непрерывность бизнеса: новый тренд или необходимость» розповідається, що неперервність бізнесу це комплексний процес, який надає компанії можливість завчасно підготуватися та визначити порядок дій для забезпечення максимально ефективного управління компанією у випадках суттєвих інцидентів та катастроф.

У публікації «Информационная безопасность – о чем говорят эксперты» розповідається про питання, що були розглянуті на конференції UA Security Conference, яку організувала та провела компанія Integrity Vision, з метою показу новітніх рішень у сфері управління інформаційною безпекою.

МИКРОГРАФИЯ И АРХИВНОЕ ХРАНЕНИЕ ДОКУМЕНТОВ

Источник: http://elar.rsvpu.ru/bitstream/123456789/16035/1/dso_2014_051.pdf
Автор: Катыхина С. А., РГППУ

Сохранение Архивного фонда, как одного из символов государственности, важнейшей части историко-культурного наследия народов и части государственных информационных ресурсов является главнейшей задачей архивной службы, приоритетным направлением работы всех архивных учреждений.[1]

Сегодня мир вновь обратился к апробированной и надежной технологи долговременного хранения информации – микрографии, существенно усовершенствованной и обогащенной новыми возможностями.

Микрография – это технология прямой репродукции документальной информации на светочувствительном пленочном материале со значительным, до 150 крат, уменьшением относительно оригинала.

Пленочный носитель называют микроформой.

Наиболее распространены 16/35 мм микрофильм, апертурные карты и микрофиши. Изображение на микроформе геометрически подобно оригиналу. Для воспроизведения микроформы требуется увеличение изображения при помощи микрографической техники.

Под микрографическими технологиями сегодня понимают не только репродукцию бумажных документов на микроформу, но весь спектр технологий для переноса бумажных и электронных документов на микроформы и обратно, а также хранения и использования документов в микрографических архивах.[2]

На сегодняшний день наиболее востребованной формой оперативного хранения информации является цифровая форма, т. е. хранение документов на магнитных лентах, магнитных дисках, магнитооптических дисках или оптических дисках. Такой архив компактен, обеспечивает скоростной доступ к информации из любой точки мира, простоту управления и поиска, одновременную работу с документом многими пользователями, очень гибкую настройку при практически неограниченном объеме хранимой информации.[3]

Цифровая форма, несмотря на ее неограниченные возможности в части оперативной работы с документами, средством долговременного хранения информации служить не может.

[1] Хабибулина Г. А. Современные проблемы создания страхового фонда копий уникальных и особо ценных документов Архивного фонда Российской Федерации. URL: http://archives.ru/reporting/report-habibilina-2012-nms.shtml.
[2] Микрография. URL: http://www.storage-systems.ru/micrography.
[3] Там же.

Как средство надежного долговременного хранения информации наилучшими возможностями обладает микрографическая форма. Микрографический архив позволяет преодолеть недостатки электронного архива, как средства долговременного хранения данных.

Хранение данных на микроформах очень консервативно, смена форматов носителей практически не происходит. Документы, перенесенные на микропленку 50 лет назад, могут быть легко воспроизведены сегодня, завтра и в будущем.

Срок гарантированного хранения микроформ составляет 100 и более лет. Микроизображение геометрически подобно изображению оригинала документа и не связано с какими-либо цифровыми форматами данных. Не требует для воспроизведения сложных устройств. При необходимости микроизображение может бать прочитано даже с помощью лупы.

Современные фотографические материалы обеспечивают высокую степень геометрического и полутонового подобия микроизображения оригиналу.

По ГОСТ 13.1.101–93 микрофильм имеет статус подлинника. Микрографический архив сегодня – это единственный путь, обеспечивающий долговременное (100 и более лет) хранение информации, в котором на уровне системного подхода решены проблемы надежности, качества и подлинности хранимой информации.

Микрографический архив – это, в первую очередь, страховой фонд документации. Однако, доступное сегодня оборудование, позволяет полностью интегрировать микрографический архив в систему документооборота современного предприятия.[2]

Сканеры микроформ позволяют любому пользователю без труда перевести в электронный вид даже очень старые документы, записанные на микропленку.

COM–технология (Computer Output Microfilm), т. е. технология вывода на микропленку цифровых данных, позволяет хранить в микрографическом архиве электронные документы, минуя бумажную форму.

COM–технология позволяет автоматически создавать образы документов, используя неформализованные данные с компьютерных систем.

COM–системы сравнивают с принтером, с тем отличием, что печать осуществляется на микрофотоноситель.

Гибридные системы представляют собой совмещенные комплекты оборудования для одновременного сканирования документов (получение электронного образа) и печати микрофильмов. Такие системы, как правило, пишут на 16/35 мм рулонный микрофильм.

Гибридные системы решают одновременно проблемы создания архивов и для оперативного и для долговременного хранения информации.

[2] Микрография. URL: http://www.storage-systems.ru/micrography.

Если проанализировать техническую сущность микрографии, нетрудно заметить, что этот процесс представляет собой сочетание фотографии и репрографии (т. е. копировальных процессов).

Типовая схема процесса микрофильмирования заключается в следующем:

– подготовка информации (документов) к микрофильмированию;

– съемка материала на специальных камерах;

– фотохимическая обработка (проявление фиксирование микропленки);

– контроль качества съемки и проявки (при неудовлетворительном качестве производится повторная съемка);

– копирование микроформ в необходимых количествах;

– укладка микроносителей в хранилище и рассылка пользователям;

– изготовление (при необходимости) бумажных копий с микрофиш;

– сканирование микроформ для передачи по техническим каналам связи и компьютерным сетям удаленному пользователю.[4]

Микрографическими архивами широко пользуются государственные структуры, государственные и коммерческие банки, национальные и публичные библиотеки, государственные архивы, научные и проектные учреждения, страховые компании, военные ведомства и т. д.

Гарантированный срок хранения информации на микрографическом носителе, без потери качества, без специальных требований к условиям хранения и при невозможности несанкционированного внесения изменений, составляет не менее 100 лет, а объемы хранения сокращаются в сотни раз.

Новые образцы оборудования значительно расширили возможности работы с микроформами, сделав их практически сопоставимыми по оперативности с электронными носителями.

В результате микрографические хранилища оказались сегодня наиболее дешевыми, надежными и удобными.

Любые данные микрографического носителя могут бать оперативно переведены в электронную форму, а данные, записанные в электронном виде, могут бать перенесены на микрографические носители, минуя бумажную форму представления.

Правительства многих стран мира законодательно утвердили подлинность документов, снятых на микрофильм, а их юридическая сила приравнена к оригиналу.

[4] Современные способы и техника создания документа. URL: http: // lib2.znate.ru / docs / index-310567. html?page=17

ФОНД ИЗДАНИЙ НА МИКРОФОРМАХ

Источник: http://www.nlr.ru/coll/ofo/micro/

Фондодержатель: Отдел фондов и обслуживания (ОФО) Российской национальной библиотеки.

Период охвата: Микроформы документов с начала письменности по настоящее время.

Объем фонда: около 500 тыс. экз. микроформ (на 01.01.2017) более 200 тыс. экз. микрофильмов; около 300 тыс. экз. микрофиш.

Состав фонда: Основной фонд структурно разделен на подфонды книг, периодических ранения на русском и иностранных языках. Частично представлены микизданий, газет, рукописных и редких изданий с учетом хронологии съемки изданий на микроформы. Он насчитывает около 500 тысяч единиц хроформы газет и диссертаций, а также изданий, которые не имеют бумажных эквивалентов, но соответствуют таким параметрам, как ценность, уникальность, высокий спрос. Подробнее о фонде.

Где можно познакомиться с изданиями фонда: Зал фонда микроформ (Московский пр., 165/2).

Кто может познакомиться с изданиями фонда: Любой читатель РНБ.

В каких эл. каталогах отражены издания фонда: Электронный каталог Генеральный алфавитный каталог книг на русском языке (1725 – 1998). Если Вы не нашли издание, нужно обратиться к консультанту по электронному каталогу в Главном здании (Зал справочной информации и электронных ресурсов), Новом здании, или адресовать запрос библиографам через форму заявки.

Шифры хранения: Шифры с префиксами Мф. Например, Мф К-1/30282

Полнотекстовые ресурсы с изданиями фонда, другие фонды, в которых представлены документы того же вида издания: Отдел рукописей, Отдел газет, Отдел нотных изданий и музыкальных звукозаписей (ОНИиМЗ), Отдел национальных литератур (ОНЛ), Отдел литературы стран Азии и Африки (ОЛСАА)

Микрофильмирование

Заказы на изготовление микрофильмов принимаются в Отделе внешнего обслуживания. Микрофильмирование производится на 35 мм негативную неперфорированную фотопленку фирм Kodak или Agfa. Максимальный размер оригинала 840х595мм.

Следует заметить, что позитив можно изготовить только с негатива. То есть позитив мы можем изготовить либо с Вашего негатива, либо с негатива, изготовленного нами, что отразится на цене.

Расценки см. в Прейскуранте на платные услуги Отдела внешнего обслуживания. Прием и выдача заказов производится в понедельник – пятницу с 10 до 17.30 (обед с 12 до 12.30)

Контактная информация:
Адрес: г. Москва, ул. Садовая, 18, отдельный вход
Станция метро: Гостиный двор
Тел.: (812) 310-98-46
E-mail: ovo-service@nlr.ru

Выписка из Прейскуранта № 7/17 на дополнительные платные услуги Российской национальной библиотеки. Стоимость услуг указана в рублях. В части услуг, облагаемых НДС по ставке 18%, сумма налога выделена. В части услуг, не облагаемых НДС, сумма НДС не выделяется.

Сканирование микрофильмов и микрофиш

Вид работ	**Ед. измерения**	**Стоимость до 300 dpi вкл. / более 300 dpi**	**В т.ч. НДС 18%**
Сканирование рулона микрофильма	1 кадр	23/40	3.51/6.10
Выборочное сканирование микрофильма	1 кадр	43/55	6.56/8.39
Сканирование микрофиш	1 кадр	20/36	3.05/5.49

Коэффициент за срочность выполнения заказа – 2.
Возможность срочного выполнения заказа заранее согласовывается между заказчиком и исполнителем.
Примечание: При сканировании копии 2-го поколения возможно ухудшение её качества.

Микрофильмирование: изготовление негативной или позитивной копии 2-го поколения

Вид работ	**Ед. измерения**	**Стоимость**	**В т.ч. НДС 18%**
При наличии в фонде РНБ микрофильма 1-го поколения.	1 кадр	23	Без НДС
При отсутствии в фонде РНБ микрофильма 1-го поколения .	1 кадр	46	Без НДС
Выборочное микрофильмирование	1 кадр	55	Без НДС

Коэффициент за срочность выполнения заказа – 2.
Возможность срочного выполнения заказа заранее согласовывается между заказчиком и исполнителем.

FAQ ПО ВОССТАНОВЛЕНИЮ ПОВРЕЖДЁННЫХ ИЗОБРАЖЕНИЙ

Источник: http://forum.ixbt.com/post.cgi?id=annc:23:28361

1. Как восстановить удалённые изображения?

1.1. При восстановлении снимков с флеш-карты рекомендуется сохранить образ карты на HDD и работать с ним, а не с картой (краткая инструкция как сохранить образ). Это ускорит поиск файлов и защитит содержимое карты от ошибок при восстановлении. Образ можно сохранить, например, с помощью HexWorkshop

1.2. Handy Recovery - универсальное восстановление удалённых файлов. Позволяет сохранять образы дисков и флеш-карт.

1.3. IsoBuster - восстановление нечитающихся файлов с CD и DVD-дисков

1.4. PhotoRescue Advanced (возможности и ссылка) - восстановление файлов с флеш-карт и минимальный ремонт

1.5. CD/DVD Inspector

1.6. PhotoRec и TestDisk - утилиты для восстановления изображений, файлов, разделов и т.д. История успеха от znznznz.

1.7. **Обсуждение восстановления информации с жёстких дисков (HDD)** - см. раздел форума "Магнитные носители информации" и FAQ по нему (в частности, пункт "Восстановление информации").

1.8. **Обсуждение восстановления информации с карт памяти (flash)** - см. раздел форума "Модули памяти" и путеводитель по нему (в частности, пункт "Обсуждение и решение проблем" и темы 1 и 2).

Восстановление удалённых файлов и отформатированных дисков с носителей информации в данной теме не обсуждается.

Для обсуждения - см. ссылки выше.

2. Файл не отображается (или отображается, но не весь), можно ли что-то сделать?

2.1. JPEGfix - утилита по оценке и ремонту JPEG-файлов. Как оценить пригодность файла к ремонту - см. пункт 3 инструкции.

2.2. Инструкция по оценке файла через упаковку в архив.

2.3. JPGscan - старая утилита по оценке пригодности JPEG-файла к ремонту.

2.4. JPEGsnoop - программа для просмотра повреждённых файлов. Краткое описание возомжностей.

2.5. ZAR, Digital image recovery - программа для поиска и извлечения изображений из пострадавших носителей, после форматирования и т.п.

2.6. Утилита Bad Peggy + другие способы автоматически проверить много файлов на годные/негодные

2.7. Почему видна маленькая картинка предварительного просмотра, а сама фотография с дефектом (не отображается)?

Маленькая картинка это эскиз/thumbnail/preview, он хранится в начале файла. Если начало файла уцелело, то эскиз/thumbnail/preview будет отображаться нормально.

2.8. Снимки в RAW рекомендую проверить в FastStone Image Viewer. Просмотр обычно показывает JPEG-preview (полноразмерный или thumbnail), опция конвертации - выгружает сами RAW-данные. Это разные блоки RAW-файла, один из них может оказаться целым или менее пострадавшим, чем другой.

3. Как можно починить файл с изображением?

3.1. JPEGfix - инструментарий по ремонту. В том числе позволяет исправить повреждения вида искажение цвета и сдвиг.

3.2. PixRecovery - онлайн ремонт и оффлайн утилита для простого ремонта файлов (JPEG, GIF, TIFF, BMP, PNG), см. также в этой теме про неё

3.3. JPEG Recovery - позволяет исправить небольшие повреждения в JPEG.

3.4. Digital image recovery 1.47 (как найти эту программу?)

3.5. JPEG Ripper от участника Dean - позволяет исправить JPEG-файл, от которого показывается только эскиз/thumbnail/preview, хотя файл большой (при условии, что основное изображение уцелело, а испорчен только заголовок).

3.6. CR2 (может быть и другие RAW-форматы) можно попробовать починить склейкой хорошего заголовка и RAW-блока из пострадавшего файла

3.7. ThumbnailExpert - программа, которая читает файлы кэша thumbnail/preview (эскизов).

4. Не справляюсь. Может ли кто-то помочь?

4.1. **Принимаю JPEG файлы на анализ ремонтопригодности**

5. Как устроены файлы с изображениями?

5.1. ITU-1150 (T.81) - спецификация на формат файла JPEG

5.2. **Очень подробная статья на habrahabr.ru об устройстве формата JPEG**

5.3. Описания сжатия и формата JPEG на CodeNet.

5.4. Сайт Independent JPEG Group - формат JPEG и исходные коды jpeglib.

5.5. Стандартные таблицы квантования и Хаффмана, если таблиц нет в изображении - подразумеваются эти.

6. Прочее

6.1. Почему при восстановлении удалённых снимков получается "каша" из других снимков?

- Из-за фрагментации файлов

6.2. Можно ли без карт-ридера восстанавливать снимки, удалённые с флеш-карты?

- Скорее всего, нет

6.3. Чем можно проверить большую коллекцию JPEG и выявить повреждённые / составить список хороших?

- ImageVerifier, обзор
- jpgtest, краткий обзор

6.4. Как проверить флешку на работоспособность?

- Отдельно от фотоаппарата (инструкция)
- Совместно с фотоаппаратом (инструкция)

6.5. Хочу использовать флешку после сбоя, что с ней нужно сделать предварительно?

- Отформатировать и проверить (см. инструкцию), если сбой повторился - не использовать.

6.6. Как отрезать лишний объем у успешно открывающихся изображений?

6.7. Как вытащить эскиз/thumbnail/preview в отдельный файл JPG, желательно автоматически - JPEG Ripper, ShowExif, exiftool

7. Вашего вопроса нет в списке?

Задавайте его в теме.

Выкладывая файлы используйте ge.tt, rghost.net, datafilehost.com или другие нормальные хостинги, где не нужно регистрироваться или смотреть рекламу и ждать чтобы скачать файл.

8. Что делать, чтобы не потерять ценные фотографии?

Главное, что нужно запомнить: **Информация - не материальна. Потерять файлы гораздо проще, чем их восстановить.**

8.1. Копирование, копирование и ещё раз копирование.

Для ценных фотоснимков рекомендую делать 3 копии: HDD компа/ноута, внешний USB-диск и DVD-диски. Делитесь снимками с друзьями, выкладывайте их в интернет. Например, яндекс-фотки - резиновые, ограничения по объёму нет, залейте туда всё в закрытые альбомы.

8.2. Проверка и ещё раз проверка. Перед удалением файлов убедитесь, что цела хотя бы ещё одна копия этих файлов. Особенно, если остаётся только одна копия. Особенно, когда используете непроверенное оборудование - в походе/отпуске/командировке. Просмотрите несколько файлов в начале, несколько - в конце и несколько - в средине - что все они скопировались, полностью и отображаются.

8.3. Перед использованием картридера убедитесь, что он поддерживает вашу флешку и её размер. Например, если картридер не знает про SDHC (4ГБ и больше), совать в него SDHC карту просто опасно, даже на чтение.

8.4. Если что-то случилось с носителем информации - не записывайте на него ничего, не удаляйте ничего, не форматируйте, не загружайтесь с него, не используйте с непроверенным оборудованием и т.д. - любая модификация уменьшает шансы спасти ценные файлы в непредсказуемое число раз. Пока вы не убедились, что спасли всё (или что

больше ничего спасти невозможно) - только читайте диск, причём чем меньше раз - тем лучше. Лучшая тактика - снять полный образ и спрятать носитель "в сейф", а данные извлекать из образа.

8.5. Проверяйте SMART ваших HDD-дисков, в т.ч. внешних USB-дисков. Reallocated_Sector отличный от нуля - повод для немедленного копирования всех данных и скорейшей замены диска.

8.6. После записи CD/DVD убедитесь, что все файлы записались успешно. Если каталогов на диске нет, достаточно проверить несколько первых, последних файлов и в средине. Если есть каталоги - то проверять нужно в корне, в первом и последнем каталоге.

8.7. Если флешка нагрелась - файлы могли скопироваться с ошибками.

8.8. Перед использованием диска больше 137 ГБ (128 GiB) на новом компе / операционной системе, **убедитесь, что комп (железо и BIOS) и ОС поддерживают такие диски** и режим поддержки включен.

8.9. Флешку для фотоаппарата - форматируйте только в нём. Воздержитесь от копирования файлов на такую флешку через картридер. Не используйте одну и ту же флешку в разных фотоаппаратах без промежуточного форматирования.

8.10. Лучше отформатировать флешку, чем стереть с неё все файлы - следующие снимки будут ложиться на флешку без фрагментирования. Кроме того, форматирование обычно быстрее.

8.11. Не стирайте файлы, если ещё хватает свободного места. Если в походе/отпуске следующий "цикл" съёмки скорее всего не заполнит флешку до конца, лучше не стирайте с флешки все файлы, хоть и скопированные на другие носители, - сотрите половину (например, менее ценные, видео, RAW при наличии JPG). Фрагментирование новых снимков - меньшее зло, чем полная потеря старых снимков из-за сбоя другой копии.

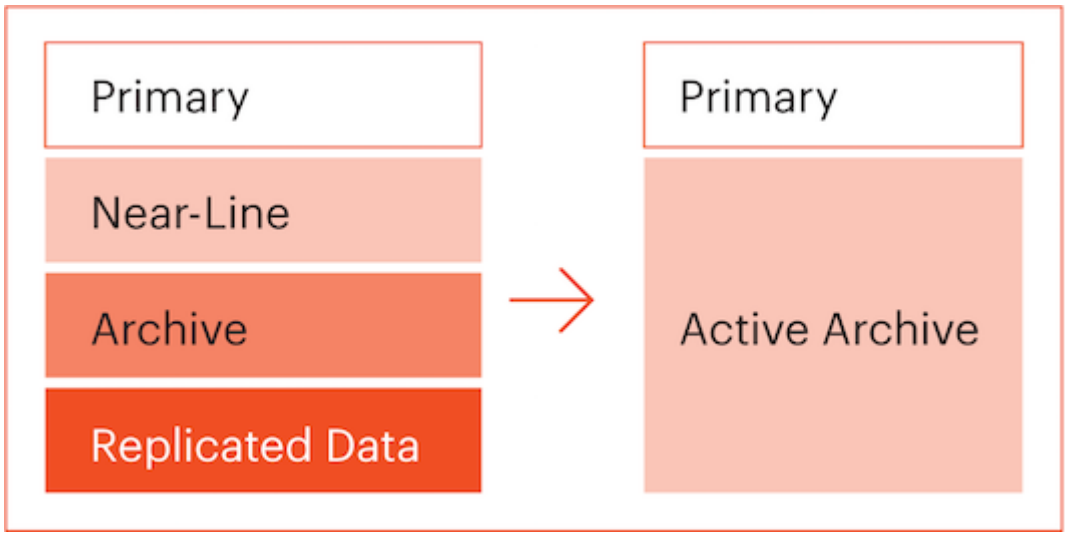
8.12. Не используйте встроенные программы Windows для просмотра фотографий - Picture and Fax Viewer, Windows Photo Viewer и т.п. Если выбора нет - **не поворачивайте фотографии в этих программах**. При повороте эти программы перезаписывают изображения, что может приводить к их искажению и потере.

АКТИВНЫЕ АРХИВЫ В ХРАНЕНИИ ДАННЫХ

Источник: http://ko.com.ua/aktivnye_arhivy_v_hranenii_dannyh_117253

С ростом объемов сохраняемой информации владельцы данных пересматривают подходы к организации хранения. Оперируя архивами в сотни терабайт, иначе смотришь на средства управления данными, их защиту, приоритеты и затраты.

Понятие «активное архивирование» вошло в оборот сравнительно недавно. Так называют долгосрочное хранение информации с обеспечением активного доступа к любой части архива, в режиме реального времени. Количество неструктурированных данных в хранилищах растет, ими надо эффективно управлять. Выделяемые в отдельный класс систем хранения, активные архивы используют индексацию, метаданные, объектную структуру и протоколы, алгоритмы защиты данных с кодом избыточности.



Повторяя иллюстрацию, можно сказать, что системы активного архива накрывают весь диапазон ответственного хранения, за исключением волатильных данных. Первичные (primary) хранилища подбираются под специфику и типы запросов критичных приложений. Все значимые данные, к которым нужен живой доступ на протяжении длительного времени, выносят в активные архивы.

Архивы – не бэкапы

Активное архивирование и резервное копирование принципиально различны: по целям, структуре данных, процедуре копирования, организации доступа.

Резервирование – это рутинное копирование операционных данных (активной и неактивной информации), для последующего восстановления работоспособности после сбоев основной системы (disaster recovery), в коротком горизонте планирования. Системы резервного копирования оптимизированы под быстрый доступ к большим объемам информации и нужны для скорого восстановления работы приложения или системы в целом.

Активные архивы объединяют наборы данных с их свойствами, детализацией и взаимосвязями (метаданными) для оперативного доступа к ним. Это актуальная информация, не копии. Причина переноса данных в архив – снижение затрат. Проиндексированный активный архив избирательно и быстро работает с индивидуальными объектами, обеспечивая им продолжительное надежное хранение.

Активное архивирование - не HSM

В отличие от систем иерархического хранения (Hierarchical Storage Management, HSM), данные активных архивов находятся в живом доступе, вне зависимости от их возраста и частоты обращения. Политики HSM построены на перемещении информации между уровнями (tiers) и устройствами хранения. Активный архив обслуживает пользователей как первичный и все остальные пулы хранения. Данные отдает то устройство, на котором они размещены. Благодаря метаданным, администрирование требует минимума времени и вычислительных ресурсов.

Архивы, активы…

Интенсивно изменяемые данные (как транзакционные базы) в системах активного архива не хранят. Чем выше объем хранения, и чем больше приложения ориентированы на чтение – тем уместнее архивирование. Данные статичной природы могут не терять ценность годами, оставаясь основным активом многих видов бизнеса. Сам термин «активы» подчеркивает важность всей хранимой информации для владельца. Разбухание архивов – сигнал для пересмотра подходов к хранению.

Размещение данных на дискретных первичных, вторичных и третичных системах хранения множит хаос пропорционально росту объемов сохраняемой информации. Сводя все более-менее статичные данные в активные архивы, пользователи получают не просто консолидированную платформу для размещения информации. Обеспечивается запас масштабирования, высокий уровень защищенности данных, быстрый поиск контента, целостность данных, мониторинг состояния систем и энергоэффективность.

Активные архивы построены на балансе цены и производительности. Скоростных стандартов для них нет, все зависит от специфики данных и приложений. Тип и количество носителей внутри архива определяются требованиями доступности данных. В целом, задержки доступа к данным систем активного архива составляют от миллисекунд до сотен миллисекунд.

Объектное хранение

Где есть временные наслоения объемных данных - появляется смысл в объектном хранении. Объектный подход удобен для размещения больших массивов неструктурированной информации: он дает свободу масштабирования, отделяет метаданные от данных, избавляет от привязки к определенной файловой системе или блочным устройствам. Администраторам незачем беспокоиться установкой уровней RAID, созданием и управлением логическими томами. Объектное хранение – естественный спутник архивирования, с запасом по росту и предрасположенностью к переносу данных в облака.

Обсуждалась проблема интеграции ленточных библиотек в объектные хранилища («вы не можете сбросить объекты на ленту»). Уже можно. Но, справедливости ради, активные архивы потому и названы так, что обеспечивают доступ к данным в реальном времени, с приемлемыми

задержками. Нужна лента – пользуйтесь пассивным архивом и средствами репликации одного в другое.

Хранилище	Блочное	Файловое	Объектное
Где?	Внутренний диск, внешний массив DAS, SAN	В файл-сервере, NAS	«в облаке»
Доступ?	К цилиндру, головке, сектору, по адресу LBA	Через файловую систему сервера	По учетной записи
Подключение?	SAS, SATA, FC, PCIe, USB, Ethernet (ПК)	Ethernet (ПК)	Ethernet (смартфон, планшет, ПК)
Протокол?	SCSI, iSCSI	CIFS, NFS, FTP	HTTP
Адресация?	# устройства, стартовый блок, смещение	F:/....	URL, http://....

RAID и Erasure Coding

Erasure coding – это подход, идущий на смену RAID в объемном хранении. Используется в системах активного архива, с разными политиками. Оригинальные данные разбиваются на фрагменты, те дополняются фрагментами с кодами избыточности, достаточно сложные алгоритмы считают и распределяют данные по носителям и серверам хранения. Запись таких фрагментированных данных требует значительных вычислительных ресурсов (потому Erasure coding применяется в архивах, менее критичных к производительности, чем к сохранности данных) . Зато архив с Erasure coding переживет отказ, скажем, 6 дисков из каждых 16. Или 6 серверов с дисками, причем они могут быть расположены на разных площадках, связанные WAN.

Актеры активного архивирования

В 2010-м году образовался альянс активного архивирования (Active Archive Alliance, AAA) – инициативное объединение для продвижения технологий продолжительного хранения с живым доступом. Участники альянса ищут пути упрощения хранения, популяризации масштабируемых решений, снижения стоимости владения, снижения рисков потери данных.

Кроме бизнеса - как суперкомпьютеры или создание медийного контента, активные архивы нужны в областях общественного интереса - как наука, образование, институты по сохранению культурного наследия. Вот, в качестве примера, детальное описание, на чем и как хранит оцифрованные архивы фестиваль джазовой музыки в Монтрё.

Обзор был бы неполным без примера программно-аппаратной реализации активного архива. Хорошо документировано решение HGST Active Archive SA-7000.

Как устроен активный архив

HGST Active Archive (далее AA) – представляет собой готовую к работе, укомплектованную систему хранения с предустановленным ПО, в формате стойки 42U. В ней есть три управляющих узла с SSD и шесть серверов хранения, к которым подключены шесть JBOD, по 98 дисков 8 TB в каждом (HGST Helium). JBOD управляются как блочные устройства хранения, по SAS. В управляющих серверах стоят SSD достаточной емкости, чтобы обслужить метаданные по более чем 1 800 000 000 объектов данных в расчете на одну стойку AA.

В AA реализовано объектное обращение к данным, по протоколу S3 или через REST API/HTTP. Физический интерфейс подключения – 6 x 10 Gb Ethernet. Для выставленных в интернет хранилищ рекомендуются фаерволы и балансировка нагрузки. В стойке есть два сетевых коммутатора – для отказоустойчивости внутренних и внешних соединений. Таких стоек на одной площадке можно разместить до шести, соединяя их интерфейсом 40 Gb Ethernet.

Программное обеспечение HGST AA разработано компанией Amplidata, которую HGST купила в 2015 году.

Производительность и возможности

Достижимая производительность для объектного хранилища по S3 – до 3.5 GB/s в операциях GET и зависит от размера объектов и количества сконфигурированных в системе потоков/демонов. Операции PUT обычно на 37% медленнее – из-за вычислений кода коррекции ошибок ECC и из-за того, что PUTs обрабатывает больше данных, чем GETs – 18 фрагментов объектов против 13, при кодировке Erasure coding 18/5.

Размер объекта	Производительность, MB/s (GET)
2MB	1580
8MB	3300
32MB	3500
64MB	3456
128MB	3328
256MB	3328

Задержка доступа к объектам не превышает <100ms в более чем 90% случаев. С добавлением стоек AA производительность масштабируется линейно. При среднем размере объекта > 1.9 MB достигается полная утилизация пространства 2.967 PB под метаданные и объектное хранение при кодировании по схеме 18/5.

Максимальный допустимый размер объекта – 16TB, а число объектов в расчете на стойку - 1.800.000.000.

Каждый контроллер стойки может обслужить 1000 активных HTTP-подключений. В стойке три контроллера, суммарной способностью 3000 подключений. При добавлении соответствующих шлюзов подключений может быть намного больше.

Защита данных

В HGST AA используется Erasure coding – подход к защите данных, превосходящий RAID в хранилищах большой емкости и масштабируемых внедрениях. Пользовательские данные непрерывно мониторятся на сбойные блоки (до 1000 однобитных ошибок подряд могут быть исправлены генерируемым системой кодом ECC). При отказе диска и потере фрагментов

некоторых объектов, все эти фрагменты восстанавливаются по оставшимся данным других дисков и распределяются по дискам по заданному алгоритму. Процесс восстановления потерянных фрагментов с помощью Erasure Coding протекает намного быстрее, чем в RAID-массиве.

HGST выбрала правило Erasure Coding BitSpread, или 18/5 – как лучший компромисс, с эффективностью 63% для объектов >512kB и высокой пропускной способностью до 3.5GB/s на стойку. Для географически распределенных реализаций HGST использует GeoSpread, или правило 18/8 – чтобы пережить полный отказ одной из площадок, еще одного JBOD и еще двух HDD, без потери данных, при 50%-й эффективности.

Каждый размещаемый объект защищен кодом коррекции ошибок ECC и разбивается на 18 фрагментов, из которых достаточно любых 13 для полного восстановления данных. Другими словами, система устойчива к потере 5 дисков. Схема 18/5 применяется ко всем объектам > 512kB.

Для объектов <= 512kB применяется политика малых объектов: схема хранения 7/5, содержащая одну копию 1:1 для ускорения доступа и копию с Erasure Coding 6/4. Сохраняется устойчивость к потере 5 дисков, как и для больших объектов.

Шлюзы

Файлы NFS или CIFS можно перемещать на HGST AA с помощью сторонних шлюзов (gateways). Если для создания объектов используются шлюзы, ими же пользуются для доступа: из-за структуры метаданных, механизмов доступа, способа разбиения объектов на блоки.

Объекты хранения на HGST AA считаются статичными - что справедливо для большинства наборов данных, переживших фазу их создания: фотографии, сканы, электронные таблицы, видеоклипы, готовые проекты, снэпшоты, бэкапы каталогов, дисков и целых систем.

База данных не может стать объектом, потому что обновляется и модифицируется слишком часто. К постоянно изменяемому контенту не обращаются как объектам - это породило бы активный пересчет ECC и перезапись всех фрагментов объектов (в нашем случае 18), бесполезную, но затратную дополнительную вычислительную нагрузку. Но пользователи вполне могут создавать базы данных объектов с привязанными к объектам метаданным.

Протоколы

S3 для облачного доступа. Система совместима с протоколом Amazon Simple Storage Service (AWS S3 - http://aws.amazon.com/documentation/s3/).

REST-API документирован в руководстве пользователя HGST AA.

Для подключения по FTP / SFTP, iRODs, NTFS и др протоколам нужны шлюзы/коннекторы с S3. Поставляются сторонними компаниями.

Управляющее ПО HGST AA – это зрелый продукт компании Amplidata. До сделки поглощения он развивался 7 лет и продавался как программно-определяемое решение хранения. С багажом HGST в создании емких дисков, JBOD и средств управления ими получившееся программно-аппаратное

решение выходит на верхний уровень в области систем хранения, оставаясь при этом привлекательным по цене.

Quanto?

Полная стойка емкостью 4.7 PB обойдется не дороже 800K евро. Можно начать со стартового набора за 300K евро. В него входят все управляющие серверы, но только один JBOD емкостью около 700 TB. Для наращивания емкости хранения докупаются JBOD с дисками.

ОПЫТ ОЦИФРОВКИ АРХИВНЫХ ДОКУМЕНТОВ В ЦЕНТРАЛЬНОМ ГОСУДАРСТВЕННОМ АРХИВЕ ГОРОДА МОСКВЫ

Источник: http://www.gdm.ru/images/infodokum_122013/tihonov.pdf

Автор: Тихонов Владимир Иванович, директор Центра автоматизированных архивных технологий Государственного бюджетного учреждения города Москвы «Центральный государственный архив города Москвы», к.и.н.

Аннотация: О практике оцифровки архивных аудиовизуальных и бумажных документов в центральных архивах Москвы и методологических проблемах долговременного хранения электронных документов.

Государственное бюджетное учреждение города Москвы «Центральный государственный архив города Москвы» (ГБУ «ЦГА Москвы») был образован в марте 2013 г. в ходе реорганизации Главного архивного управления города Москвы (Главархива Москвы). В его состав вошла большая часть неуправленческих подразделений Главархива Москвы: все центральные архивы Москвы, хранящие документы постоянного срока хранения (более 10 млн ед. хр.) и преобразованные в центры хранения документов, а также специализированные функциональные центры.

Основная цель работы государственных архивов заключается в обеспечении граждан, общества и органов власти ретроспективной документированной информацией. Одним из эффективных способов, повышающих оперативность и качество обслуживания пользователей архивными документами, является использование в работе архивов информационно-коммуникационных технологий. Их стремительное развитие в последнее десятилетие открыло для архивной отрасли некоторые совершенно неожиданные перспективы. Сканирование и оцифровка документов оказались вдруг наиболее технологичными способами копирования и распространения архивной информации, а включение электронных образов документов в информационно-поисковые системы превратило их в наиболее продвинутые средства работы с фондом пользования архивными материалами. В прошедшие несколько лет в России

не осталось, наверное, ни одного архива, который не попытался бы поставить технологии оцифровки на пользу обществу и государству.

В архивной отрасли Москвы первые опыты по оцифровке документов и созданию электронного фонда пользования начались в Центральном архиве аудиовизуальных документов Москвы (ЦААДМ).[1] Это был естественный и закономерный процесс, так как оцифровка была признана практически единственной технологией, направленной на обеспечение сохранности аудиовизуальной информации и создания копий их страхового фонда.

С 2004 г. в ЦААДМ проводят оцифровку фонодокументов, хранящихся на магнитной ленте. В настоящее время оцифровано уже более 1100 ед. хр. Из общего фонда в 4629 ед. хр. (23,5 %) В качестве кодека используется PCM (Pulse Code Modulation), преобразующий звуковые сигналы в цифровой формат несжатого (т.е. без потери качества) звука. Оцифровка проводится с частотой дискретизации 44,1 КГц и разрядностью 16 бит. До 2011 г. запись фонда пользования проводилась на оптические диски CD. В настоящее время используются диски DVD, причем, каждый фонодокумент записывается на два отдельных диска.

Планомерную оцифровку фотодокументов (негативов, позитивов и слайдов) в ЦААДМ начали с 2005 г. Сканирование производится с разрешением 2000 точек на дюйм в страховом формате TIFF, после чего создается дополнительная копия в пользовательском формате JPEG. К концу 2013 г. оцифровано более 13 тыс. фотодокументов, что составляет 5 % всего фонда архива. Файлы хранятся в трех экземплярах: на двух отдельных оптических дисках (TIFF-файлы на DVD, JPEG-файлы на CD), а также на файл-сервере, интегрированном с автоматизированной информационно-поисковой системой (АИПС) архива. Таким образом, пользователь системы может не только искать фотодокументы по атрибутам описания, но и просматривать их электронные образы.

Но наиболее продвинутые технологии оцифровки и электронного фонда пользования нашли применение в так называемом Видеоархивном комплексе Главархива Москвы, который был создан в соответствии с постановлением Правительства Москвы от 18.01.2005 No 27-ПП и распоряжением Правительства Москвы от 18.07.2006 №о 1381-РП. В 2007 – 2008 гг. велась подготовка и реализация проекта, опытная эксплуатация оборудования.

С января 2009 г. Видеоархивный комплекс был введен в промышленную эксплуатацию, и на тот момент являлся, пожалуй, самым большим и самым сложным хранилищем цифровой аудиовизуальной информации в архивных учреждениях не только России, но и Европы.

Он оснащен мощными серверами, дисковыми накопителями и ленточными библиотеками производства IBM, которые вмещают 1927 картриджей с лентами LTO-3 совокупной емкостью в 750 терабайт.

[1] В настоящее время – это Центр хранения электронных и аудиовизуальных документов Москвы (ЦХЭиАДМ) ГБУ «ЦГА Москвы».

Это позволит пользователям получать оперативный доступ к более чем 10 тыс. часов видеодокументов. Катастрофоустойчивость комплекса обеспечивается размещением однотипного оборудования на разных территориях Главархива Москвы, которые объединены выделенным каналом связи в 100 Мбит/с. Средства четырех видеомонтажных станций способны конвертировать, импортировать и экспортировать видеодок дументы в самые разнообразные цифровые и аналоговые форматы. Но основными форматами хранения в ленточной библиотеке являются форматы, обеспечивающие компрессию видеосигнала без потерь информации по стандартам MPEG-2 50Mbit I-frame 4:2:2 и 1080I MPEG2 50Mbit Long GOP 4:2:2 (с битрейтом до 50 Мбит/с), а также формат файлового контейнера MXF OP1A. Помимо размещения в ленточной библиотеке, видеодокументы, после соответствующего преобразования, записывают на внешние носители (DVcam и DVD), которые передаются в традиционное архивохранилище и рассматриваются в качестве страхового фонда.

Таким образом, оцифровка видеодокументов ЦХДЭ и АДМ преследует несколько целей. Во-первых, это – создание страховых копий в цифровых форматах, так как состояние видеопленок Betacam и VHS уже давно вызывает большие опасения. Во-вторых, создание автоматизированного фонда пользования. АИПС Видеоархивного комплекса позволяет вести расширенный поиск видеодокументов по любым атрибутам, но главное, с ее помощью на любое рабочее место в Главархиве Москвы можно «вызвать» и просмотреть цифровые копии видеодокументов в формате MPEG-2 низкого разрешения (1 Мбит/с). В-третьих, оборудование комплекса и хранящиеся в ленточной библиотеке полные цифровые копии видеодокументов предназначены для создания их копий по запросам пользователей без дополнительного использования оригиналов, что должно благотворно сказаться на обеспечении их сохранности. При этом, по желанию пользователя возможно проведение монтажа видеороликов из фрагментов разных документов, а также реставрация изображения и звука. Наконец, в Видеоархивный комплекс загружают новые поступления видеодокументов в архив, причем эти поступления – уже изначально в цифровых форматах. Так, что ленточные библиотеки используются, в том числе, как временные хранилища для документов, до проведения более тщательной экспертизы их ценности и последующего отбора на постоянное хранение. Остается добавить, что в настоящее время оцифровано и загружено в Видеоархивный комплекс более 200 часов архивных видеодокументов (15,1 % стоящих на учете) и несколько сотен часов новых поступлений.

Не менее насущной для ГБУ «ЦГА Москвы» является задача оцифровки и создания электронного фонда пользования архивными документами на бумажной основе. В 2008 г. в соответствии с п. 3.5.1 «Плана мероприятий по проектированию, разработке и внедрению информационных систем на 2008 год», утвержденного распоряжением Правительства Москвы от 11.04.2008 No 753-РП, в нескольких десятках органов исполнительной

власти Москвы предстояло реализовывать проект по созданию Единой системы электронных архивов документов Правительства Москвы. Ядром проекта являлось сканирование архивных документов и создание на этой основе распределенного хранилища их электронных копий. В августе 2008 г. наступила очередь московских архивов предоставить исполнителю по государственному контракту архивные документы для перевода их в оцифрованную форму.

Известно, что оцифровка архивных материалов – процесс весьма трудоемкий, а потому дорогостоящий. С самого начала проекта было ясно, что перевести в цифровой вид удастся лишь несколько тысяч архивных дел. Поэтому перед московскими архивистами встали две взаимосвязанные задачи: определить состав документов, подлежащих оцифровке, и способ использования подготавливаемого информационного ресурса. При решении первой задачи выбор был остановлен на фондах райисполкомов Москвы (ЦАГМ)[2], документы которых за 1950 – 1980-е годы имеют важное социальное значение и по которым ежегодно исполняется почти треть из 60 тыс. социально-правовых запросов, поступающих в службу «одного окна» Главархива Москвы.

Решение второй задачи виделось во включении электронных копий документов в подсистему «Научно-справочный аппарата» Интегрированной автоматизированной информационной системы (ИАИС) Главархива Москвы. Это означало не только модернизацию системы, но и проведение индексации файлов с оцифрованными образами. В систему должны были поступать не только архивные шифры, номера, даты и заголовки документов, но и сведения обо всех персоналиях и адресах зданий, которые встречались в текстах решений райисполкомов.

Забегая вперед, стоит отметить, что такой подход увеличил трудозатраты при создании информационного ресурса, однако, его эффективность сразу же сказалась на существенном повышении оперативности исполнения запросов и на улучшении ситуации с обеспечением сохранности используемых архивных документов.

В настоящее время приказом Главархива Москвы, при возможности выводить копии документов из ИАИС, запрещено выдавать дела из архивохранилищ ЦХД после 1917 г.

Оцифровка документов ЦАГМ проводилась на территории архива, но на оборудовании и специалистами исполнителя по государственному контракту. Для этого в здании архива были оборудованы 10 рабочих мест с планетарными сканерами (9 сканеров для оцифровки документов форматом до А2 и один сканер – для формата А1). Кроме того, были организованы еще 6 рабочих мест для расшивки и последующей сброшуровки сканированных архивных дел. Тестовое сканирование архивных листов на сканерах с автоматической протяжкой листов показало, что данная технология негативно влияет на физическую сохранность архивных документов.

[2] В настоящее время – Центр хранения документов после 1917 г.

В связи с этим от использования поточных сканеров отказались с самого начала. Сканирование осуществлялось с разрешением в 300 точек на дюйм, с сохранением результирующих файлов в форме TIFF. Выборочный контроль качества оцифровки проводили как работники исполнителя, так и архивисты. При хорошем качестве цифруемого материала сканирование проводилось в «бинарном режиме». Однако в половине случаев пожелтевшая бумага документов и нечеткость машинописи потребовали проводить сканирование в «градациях серого», что значительно увеличивало объем файлов с оцифрованными образами документов. Так что, в зависимости от исходных документов (формат которых мог достигать размера А0 и больше), объем файла мог достигать 50 - 60 Мбайт.

По условиям государственного контракта, оцифрованный материал исполнитель перевозил на свою территорию, где силами собственных сотрудников осуществлялась его индексация.

Описание оцифрованных образов проводилось на уровне каждого документа (распоряжения райисполкома) и по согласованным с Главархивом Москвы правилам, основанным на «Методических рекомендациях по описанию и классификации документной информации для создания АИПС "Организационно-распорядительные документы Московского городского совета и Исполнительного совета народных депутатов за 1931 – 1991 гг."» (Мосгорархив, 2003).

Собственно оцифровка архивных документов заняла около четырех месяцев. А вот этап индексации оказался наиболее продолжительным и растянулся почти на полгода. Результаты многомесячной работы передавались Главархиву Москвы постепенно, по мере готовности. Данный информационный ресурс представлял собой электронные копии документов, записанные на диски DVD в двух экземплярах. Каждая партия переданных DVD-дисков сопровождалась CD-дисками, содержащими базу данных с индексами и описанием архивных документов. Передача осуществлялась по описям дисков с электронными копиями архивных документов, которые включали следующие статьи описания:

– номер диска по описи;
– идентификация, шифр диска (включавший номер описи, номер диска по описи и обозначение его экземплярности с рабочими или резервными копиями архивных документов);
– номера архивного фонда, описи, архивных дел, копии документов которых записаны на диск);
– количество оцифрованных дел, копии документов которых записаны на диск;
– количество электронных копий архивных документов, записанных на диск;
– количество компьютерных папок, составляющих электронный массив;
– количество компьютерных файлов, составляющих электронных массив;
– общий объем файлов в мегабайтах и байтах;
– дата записи электронного массива на диск.

В итоговую запись к описи вносилось количество: дисков, оцифрованных дел, электронных копий листов архивных документов, компьютерных папок и файлов.

После приема носителей с оцифрованными образами документов наступал следующий этап работы по подготовке информационного ресурса к использованию. В Информационном центре Главархива Москвы с помощью специализированного программного проводилась проверка технического состояния всех поступивших дисков.

Практика показала, что на проверку одного DVD-диска с записью нескольких сотен графических файлов необходимо 30 - 40 минут. В результате многомесячной работы было выявлено около 200 дисков (2%), оказавшихся в неудовлетворительном состоянии, которые исполнителю пришлось заменить. При этом нередко отмечалась выбраковка до десятка дисков подряд (принадлежащих к одной серийной партии), что подтверждает одно из базовых правил подготовки электронных документов к долговременному хранению – запись рабочих и резервных экземпляров документов на электронные носители разных фирм-производителей.

Завершающим этапом формирования архивного информационного ресурса был импорт электронных копий документов и их метаданных (индексной информации) в ИАИС Главархива Москвы. На этой стадии стандартными средствами СУБД выявлялись и исправлялись ошибки в индексации и в привязке оцифрованных образов к архивному шифру документов.

Окончательное мнение о полноте и качестве электронной картотеки по документами райисполкомов Москвы должны были высказать ее конечные пользователи – работники ЦАГМ и управлений Главархива Москвы, которые приступили к ее эксплуатации и, в целом, остались довольны ее поисковыми возможностями.

В 2011 – 2012 гг. работы по оцифровке архивных документов были продолжены. На этот раз, кроме райисполкомов выбор пал на фонд Мосгорисполкома (Моссовета): на дела, содержащие постановления, распоряжения, протоколы и другие распорядительные и нормативные документы. В итоге в 2008, 2011 и 2012 гг. были оцифрованы более 1,8 млн документов, входящих в состав 8126 дел по фондам шести райисполкомов (из 30), совокупным объемом более 1,6 млн страниц. В картотеку райисполкомов были включены около 17 млн фамилий и 20 млн географических названий. То обстоятельство, что к 2011 г. в архиве была полностью завершена электронная картотека «Решения и распоряжения Мосгорисполкома», сократило объемы работ по индексации документов Мосгорисполкома и позволило довести этот оцифрованный массив до 8712 дел (более 2,5 млн страниц).

С 2011 г. изменились форматы записи оцифрованных образов и электронные носители, на которых они принимались от исполнителя. Теперь использовался формат JPEG и внешние жесткие диски большой емкости. В

настоящее время файлы с электронными копиями документов ЦХД после 1917 г. хранятся в Центре автоматизированных архивных технологий ГБУ «ЦГА Москвы» в трех экземплярах: в двух экземплярах на 2792 дисках DVD (райисполкомы) и 24 внешних жестких дисках (Мосгорисполком и райисполкомы) и один экземпляр на 238 дисках BD-R в роботизированном дисковом накопителе, интегрированном с ИАИС. Совокупный объем сохраняемых оцифрованных образов, таким образом, составляет почти 25 Тбайт данных.

Конечно, 17 тыс. оцифрованных дел – это слишком незначительный процент от всего объема Архивного фонда города Москвы. По нашим оценкам, при сохранении существующих темпов оцифровки (читай, объемах финансирования этих работ в 2008 – 2012 гг.) потребуется не менее сотни лет на то, чтобы перевести в электронную среду лишь наиболее востребуемую часть фондов Центрального государственного архива города Москвы. Учитывая острую потребность в проведении данных работ, Правительство Москвы в конце 2012 г. поставило для московских архивов целую линейку современного сканирующего оборудования вместе с емкими серверами и системами хранения информации. В середине 2013 г. ГБУ «ЦГА Москвы» организовало собственный производственный участок и приступило к обучению сотрудников и опытным работам по оцифровке документов.

Еще одним объектом для оцифровки в 2013 г. стали наиболее востребованные, но отсутствующие в свободном доступе в читальном зале, описи ЦХД до 1917 г.[3]

За полгода Центр микрографии и реставрации документов ГБУ «ЦГА Москвы» сканировал почти 700 описей по 340 фондам. Для использования этого информационного ресурса Центром автоматизированных архивных технологий была разработана и внедрена в эксплуатацию поисковая система, в которой посетители читального зала могут получить доступ и просмотреть образы 40 тыс. листов описей.

В ближайший год этот электронный фонд пользования описями составит более 2 тыс. архивных справочников только по ЦХД до 1917 г., после чего будет пополняться электронными копиями описей других центров хранения документов архива.

В заключение следует отметить, что оцифровка архивных документов – это только начальный этап создания полноценных систем доступа к архивной информации. Одновременно с этим необходимо решать не менее сложную и дорогостоящую задачу сохранения оцифрованного контента. И проблема здесь не только в выборе надежных и емких систем хранения сотен терабайт данных (выбор аппаратных и программных компонентов, электронных носителей), но и в последующих, неоднократных переводах информационно-поисковых систем на новые технологические платформы.

[3] До апреля 2013 г. – Центральный исторический архив города Москвы (ЦИАМ).

Не только практика, но методология хранения электронных (оцифрованных) документов в нашей стране находится в зачаточном состоянии.

Цели и задачи электронных архивов, создаваемых в организациях и органах власти, определяются, как правило, на ближайшую перспективу, без учета требований долговременного хранения документов.

Да, и эти требования никто не торопится формулировать, не говоря уже об их нормативном оформлении. Такой подход уже в недалеком будущем, через 15 - 20 лет, грозит практически полным исчезновением баз данных с документами, вышедшими из оперативного использования. Повторится судьба электронных информационных ресурсов 80 – 90-х годов прошлого века. Те же немногие ценные базы данных, в том числе с оцифрованными архивными документами, ждет дорогостоящий перевод на более современные платформы. Однако издержки на проведение миграции данных можно было бы снизить, если уже сегодня при создании таких информационных систем ориентироваться на широко распространенное программное обеспечение и внутреннюю организацию баз данных, облегчающую экспорт документов.

Таким образом, без стандартизации и унификации принципов и подходов к созданию информационных систем с ценными электронными документами, на государственном уровне или на уровне профессиональных ассоциаций, будет практически невозможно обеспечить сколь-нибудь длительное сохранение цифрового богатства современной России.

КАК ПОСТРОИТЬ ЭФФЕКТИВНЫЙ ЭЛЕКТРОННЫЙ АРХИВ НА ВАШЕМ ПРЕДПРИЯТИИ

Источник: http://efsol.ru/articles/create-an-electronic-archive.html

Много предприятий сегодня используют электронный архив документов (ЭА). Но каждая из них вкладывает свой смысл в это понятие. Некоторые понимают это как место на диске, где просто хранятся скан-копии оригиналов документов. В отдельных случаях – скан-копии документов, которые привязаны к документам учетной системы. Но в нашем понимании электронный архив – это нечто большее, чем просто хранилище изображений.

Эффект от использования системы ЭА напрямую зависит от того, по какому принципу построен архив и насколько удобно осуществляется взаимодействие пользователя с системой.

В этой статье мы рассмотрим несколько самых распространенных технологий создания и ведения электронного архива и сделаем вывод о том, какая же из них приносит наибольший эффект для бизнеса.

Стоит сделать оговорку: мы не говорим об эффективности в целом, а только об одном из параметров эффективности автоматизированных систем. Когда речь идет об автоматизации бизнес-процессов, то насколько эффективна система, можно судить по количеству выполняемых вручную операций. Очевидно, что чем меньше таких операций, тем система более продуктивна и дает больший бизнес-эффект.

А разве хранилище сканов — не электронный архив?

Бывают случаи, когда электронный архив строится по принципу обычного бумажного архива. Отличие лишь одно – документы разбросаны по папкам и лежат мертвым грузом не в какой-то комнате, а на электронном носителе. Работать с таким архивом, конечно же, удобнее, чем с оригиналами, но лишь не намного.

Сотрудникам при обработке документов, по-прежнему, надо вводить их вручную в учетную систему, а при поиске документов задаваться вопросами: «А тот ли это документ, что я ищу? А в актуальной ли он редакции?»

Но все же окончательно сказать о том, что такой архив не имеет права на жизнь, мы не можем. Безусловно, эффект от работы будет, но только в маленьких компаниях, где документооборот не превышает 100 в месяц. Сравним количество ручных операций и автоматизированных на примере пачки входящих документов.

Таблица 1 – Сравнение количества ручных и автоматических операций при работе с ЭА

Ручные операции	**Автоматизированные операции**
1. Сортировка документов, разделение на пакеты. 2. Сканирование документов. 3. Переименовать скан-копии документов. 4. Рассортировать документы по папкам. 5. Создать документ в учетной системе. 6. Вручную ввести данные со скан-копии. 7. Провести документ.	1. Поиск документа в электронном хранилище.

В таких условиях практически все операции так и остаются ручными. Это отнимает много времени, снижая эффективность работы с архивом. В статье «Мифы о системах потокового ввода» вы можете посмотреть, сколько времени отнимает ручной и автоматизированный ввод документов в учетную систему.

На уровень выше – сканы имеют связку с документами (бухгалтерскими проводками) учетной системы

Некоторые компании пошли дальше. Тут из транзакций учетной системы можно открыть скан-копию оригинала документа. В таком случае эффективность работы повышается, но много работы приходится, как и раньше, выполнять в ручном режиме, совершать много лишних действий.

Таблица 2 – Сравнение количества ручных и автоматических операций при работе с ЭА

Ручные операции	**Автоматизированные операции**
1. Сортировка документов, разделение на пакеты. 2. Сканирование документов. 3. Переименовать скан-копии документов. 4. Рассортировать документы по папкам. 5. Создать документ в учетной системе. 6. Вручную ввести данные со скан-копии. 7. Прикрепить скан-копию к документу учетной системы. 8. Провести документ.	1. Быстрый поиск документа в учетной системе. 2. Открыть скан-копию из транзакции учетной системы.

Как и в предыдущем случае, большинство операций производится вручную. Небольшое преимущество появляется лишь на этапе поиска электронной скан-копии. Найти документ в учетной системе – задача несложная, а уже потом открыть скан из транзакции – не занимает много времени.

Также стоит отметить тот факт, что в таких условиях работы с архивом, операции строятся последовательно. То есть только закончив работу с одним документом, можно переходить к следующему и выполнять те же действия.

Эффект от такого электронного архива выше, чем в предыдущем случае, но не настолько высокий, как мог бы быть.

Электронный архив как часть системы электронного документооборота

В погоне за повышением эффективности работы с электронными копиями документов, за автоматизацией документооборота были созданы комплексные решения – системы электронного документооборота (СЭД). ЭА представляет собой некую подсистему, часть СЭД, интегрированную с учетной системой. Такой подход к автоматизации архивов качественно отличается от двух предыдущих.

Функциональные возможности таких решений очень широки, поэтому описывать их мы не будем. Вместо этого мы приведем перечень основных задач, которые ЭА призван решить:

- Оперативное реагирование на запросы государственных уполномоченных органов с требованием предоставить документы (налоговые органы, внебюджетные фонды и т.д.), а также внутренним контролерам.
- Контроль своевременного предоставления контрагентами входящей финансово-первичной документации (ФПД) и правильности ее оформления.
- Снижение востребованности бумажных оригиналов с возможностью вывоза бумажного архива в удаленное защищенное хранилище и др.

Эффективность ЭА гораздо выше в сравнении с остальными методами ведения электронного архива. Поэтому мы подробнее опишем данную технологию и расскажем, почему же работать с автоматизированными системами действительно продуктивно.

В первую очередь отметим, что электронный архив подразумевает под собой не только хранение документов, но и наполнение этого архива. На рисунке 1 изображена схема автоматизированного ввода документов в архив.

Рис. 1 – Процесс ввода документов в ЭА

На сканирование подается пачка документов, которые можно не сортировать. Со сканера документы отправляются в интеллектуальную систему распознавания, которая сама определяет вид документа, его атрибуты и сама знает, где закончился один документ и начался следующий.

На каждом предприятии процесс работы с документами имеет свои особенности. И даже на одном предприятии может существовать несколько сценариев работы с одним и тем же документом. Поэтому мы рассмотрим только один из возможных сценариев, чтобы продемонстрировать какое место электронный архив занимает на участке документооборота.

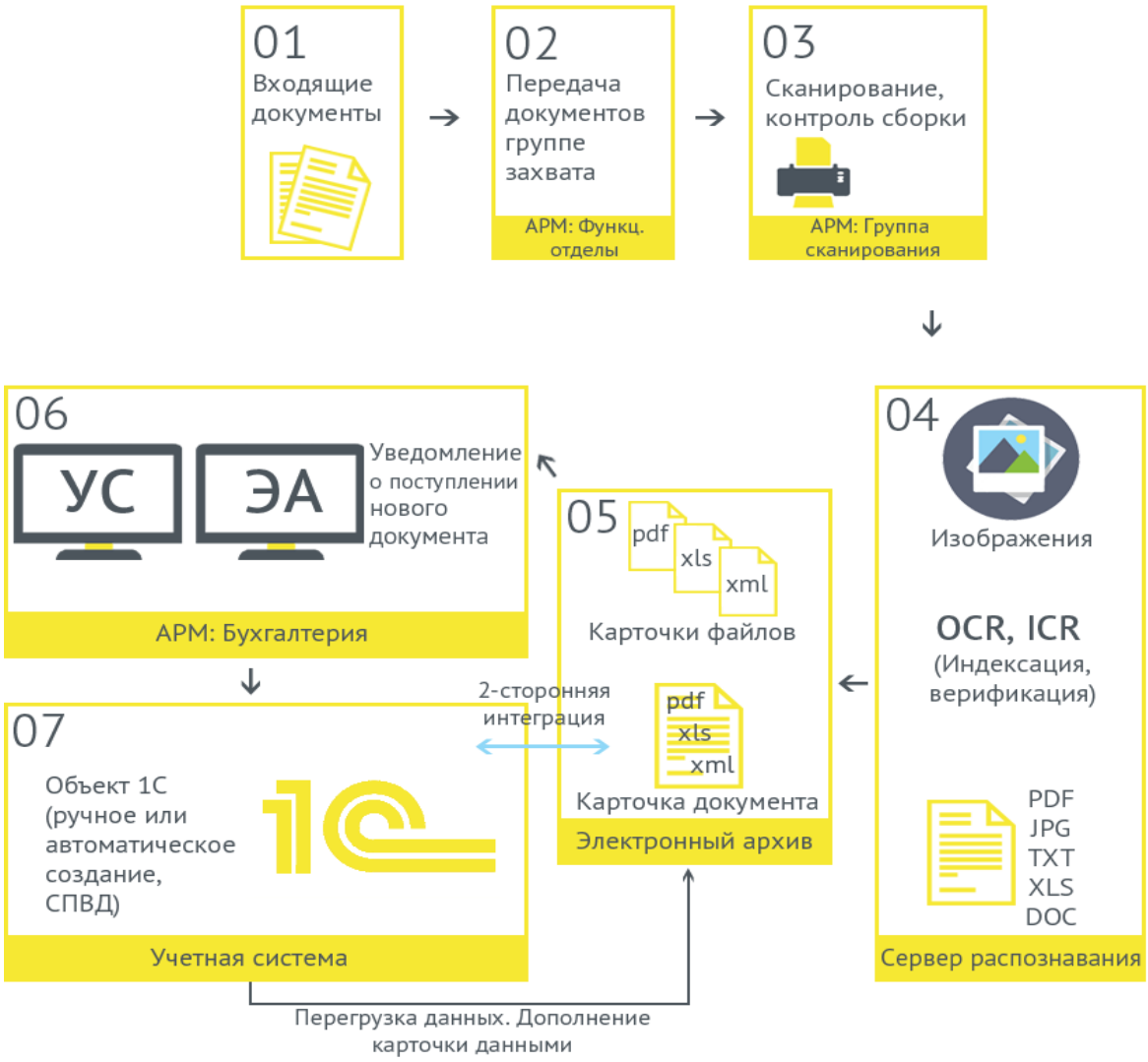


Рис. 2 — Сценарий работы со входящими документами

Аббревиатуры:

1. *СПВД — система потокового ввода документов*
2. *АРМ — автоматизированное рабочее место*
3. *УС — учетная система*

Опишем приведенный сценарий поэтапно:

01	Поступают входящие/исходящие первичные документы.
02	Функциональные отделы принимают товары и документы. Передают документы сотруднику, ответственному за сканирование.
03	ФПД сканируются в едином потоке.

04	Система OCR/Data Capture автоматически ориентирует страницы, находит начало документа и конец (автоматическое разделение на документы), распознает, извлекает атрибуты, формирует изображение с оптимальными соотношением размер файла/качество, формирует структурированные данные из этого документа (вид документа, номер, дата, реквизиты контрагента/организации, табличная часть: номенклатура, единицы измерения, цена, количество, сумма, НДС; и т.д.)
05	Данные передаются в систему ЭА, автоматически создается атрибутированная карточка, прикрепляется скан-образ.
06	Бухгалтер получает уведомление о поступлении нового документа.
07	Бухгалтер заходит в систему ЭА и при нажатии кнопки, структурированные данные документа перегружаются в учетную систему, где автоматически создается нужный документ УС, готовый к проведению.

Количество ручных операций сокращено до минимума. Все процессы выполняются автоматически, оставляя за сотрудниками лишь функцию контроля.

Как мы уже и говорили, снижая количество ручных операций, повышается эффективность работы с электронным архивом.

Системы потокового ввода приносят наибольший бизнес-эффект для предприятия при работе с электронными копиями документов.

В чем же кроется эффективность электронных архивов

Какую же систему можно назвать самой эффективной? Эффективной системой можно считать ту, которая будет решать задачи вашего бизнеса, сокращая всевозможные убытки. Для предприятий с малым документооборотом достаточно будет и простого структурированного хранилища сканов. При этом такой архив будет приносить свой бизнес-эффект и решать поставленные задачи. Но если у вас требования шире, чем просто хранить электронные копии документов, то нужно задумываться о внедрении комплексного решения, которое позволит максимально автоматизировать документооборот предприятия. ЭА как комплексное решение просто необходимо компаниям с большим документооборотом, с территориально распределенными филиалами, предприятиям, которым приходиться часто сдавать налоговые и другие отчетности.

НЕПРЕРЫВНОСТЬ БИЗНЕСА: НОВЫЙ ТРЕНД ИЛИ НЕОБХОДИМОСТЬ

Источник:
http://ko.com.ua/nepreryvnost_biznesa_novyj_trend_ili_neobhodimost_120497
Автор – Виктория Борсуковская

В последнее время все больше компаний обращаются к вопросу комплексной защиты своих ресурсов, обеспечения полноценного функционирования внутренней инфраструктуры и минимизации рисков утери важных данных, именно это и принято называть непрерывностью бизнеса.

Непрерывность бизнеса это комплексный процесс, который помогает определить потенциальные угрозы и их влияние на обычный порядок ведения бизнеса. Указанный процесс предоставляет компании возможность заблаговременно подготовится и определить порядок действий для обеспечения максимально эффективного управления компанией в случаях существенных инцидентов или катастроф.

Сегодня, практически все участники рынка стремятся к наибольшей автоматизации бизнес-процессов и уменьшению человеческого влияния на их внедрение. С одной стороны это положительное стремление. Тем не менее, с другой стороны – принимая во внимание увеличивающееся количество и сложность кибератак – автоматизация становится наиболее уязвимым элементом компании, который может на значительный период времени остановить или парализовать ее деятельность.

В своих исследованиях Gartner указывает, что менее 30 % компаний, которые входят в рейтинг Fortune 2000, выделяют финансовые ресурсы на развитие и внедрение процессов управления непрерывностью бизнеса. Низкий уровень заинтересованности компаний в выделении отдельного процесса и инвестировании в его развитие может быть вызван сложностями в понимании технических процессов. Плюс к этому достаточно высокой стоимостью ИТ-продуктов, которые обеспечат внедрение автоматизированного процесса бизнес непрерывности компании.

Одним из базовых международных стандартов для управления непрерывностью бизнеса является ISO 22301, основной задачей которого является направление компании в поиске эффективного подхода к минимизации последствий, вызванных рисками и угрозами, которым она подвергается.

В соответствии с ISO 22301 непрерывность бизнеса это комплексный процесс управления, который определяет потенциальные угрозы для компании и их влияние на функционирование бизнес процессов, а также определяет основы для создания эффективной системы, способной сдерживать и противодействовать таким угрозам.

Непосредственно сам процесс управления непрерывностью бизнеса очень прост схематически. Тем не менее, он содержит множество элементов и ключевых вопросов, которые должны быть приняты во внимание при разработке системы непрерывности определенной компании.

Рис. 1. Жизненный цикл процесса управления непрерывностью бизнеса

Довольно много компаний считают, что непрерывность бизнеса это исключительно ИТ-процесс. Обычно, термины непрерывность бизнеса (business continuity) и аварийное восстановление (disaster recovery) используются как синонимы. Эти понятия являются взаимосвязанными, но выполняют разные функции внутри организации.

Все компании, кто раньше, кто позже, сталкиваются с вопросом определения ответственности за непрерывность бизнеса и аварийное восстановление. Федеральный Совет по вопросам надзора за финансовыми учреждениями США (FFIEC) определяет, что правление и топ-менеджмент компании обязаны обеспечить процесс определения, оценки, приоритезации, управления и контроля над рисками, как части процесса планирования непрерывности бизнеса.

Это означает, что правление и топ-менеджмент компании являются непосредственными собственниками программы непрерывности бизнеса и несут ответственность за разработку и наличие соответствующих регуляторных документов, а также осуществление контроля над их надлежащим исполнением. В то время, когда высшее руководство осуществляет контроль и надзор над исполнением документов (рис.2), менеджеры и персонал несут ответственность за внедрение и обновление соответствующих планов по обеспечению непрерывности бизнеса.

Это значит, что к процессу непрерывности привлекается весь персонал – как бизнес-функция, так и технические специалисты. Именно

поэтому непрерывность бизнеса не может рассматриваться исключительно как ИТ-функция. Без привлечения представителей бизнес-функции, эффективность планов непрерывности может быть значительно ослаблена, а время, отведенное на аварийное восстановление, может быть достаточно длительным или использовано не оптимально.

Рис. 2. Пример построения контроля в организационной структуре компании

Также, структура непрерывности бизнеса должна включать создание команды кризисного управления (рис. 3), которая состоит из первых лиц и представителей высшего руководства; т.е. лиц, которые обладают достаточными знаниями о работе и функционировании компании, и уполномочены принимать критические решения в кризисных ситуациях. Эта команда должна иметь документальный план, который определяет руководителя кризисного комитета, его координаторов и потенциальных заместителей, а также определяет порядок информирования, коммуникации и процесс функционирования во время кризисной ситуации и после ее ликвидации.

Команда осуществляет надзор и координирует процесс взаимодействия с командами аварийного восстановления (техническими и бизнеса), определяет лиц, которые уполномочены взаимодействовать с внешними участниками до, во время и после инцидента/кризиса/катастрофы. Определенные дополнительные функции могут предусматривать контроль и управление вопросами охраны жизни и здоровья сотрудников и связанных лиц, своевременной и четкой оценки инцидента и его последствий, доступности персонала и ресурсов, необходимых для восстановления, минимизации имущественных и финансовых убытков, и др.

В Украине, в большинстве случаев, вопрос непрерывности бизнеса важен для компаний с иностранным капиталом, которые обязаны выполнять как национальные стандарты, так и следовать стандартам, установленным на групповом уровне. Для внедрения процесса управления непрерывностью бизнеса могут применяться разные модели**,** в частности:

1. Внешние поставщики услуг – консалтинговые компании, которые разрабатывают и предоставляют решение под ключ для конкретной компании.

2. Внутренние ресурсы – определение специалиста или отдельной службы, которые обеспечат анализ и внедрение решения на месте.

3. Смешанный вариант – первичный анализ проводится внешней компанией, а непосредственно внедрение проекта обеспечивается силами самой компании.

Рис. 3. Организационная структура управления непрерывностью бизнеса

Этапы внедрения процесса непрерывности бизнеса:

I. Разработка базовых нормативных документов, регулирующих соответствующий вопрос.

II. Проведение анализа влияния кризисных ситуаций/инцидентов на деятельность и процессы компании.

Составляется полная картина деятельности компании, формируется перечень процессов/функций, а также определяется тип влияния на бизнес (материальный, экономический, репутационный), зависимость от информационных ресурсов и максимальное время простоя.

III. Анализ и оценка рисков.

Такая деятельность обеспечивает понимание и принятие потенциальных угроз, а также их последствий и источников, определяет уязвимые места компании и позволяет найти возможные варианты действий, чтобы избежать или устранить такие угрозы или последствия.

Оценка рисков является основой для дальнейшей разработки стратегии непрерывности бизнеса и оптимальных сценариев ее реализации.

IV. Разработка стратегии компании.

Стратегии включат модельные сценарии, последствия которых негативно влияют на компанию, блокируют или ограничивают ее деятельность. Обычно, такие сценарии охватывают вопросы недоступности помещения, персонала, ограниченное использование или утрату ИТ-ресурса, и др. Сценарии дают возможность объективно рассмотреть работу компании, определить приоритетные направления/процессы, минимально необходимые ресурсы для восстановления, а также меры для минимизации рисков возникновения таких сценариев.

V. Разработка и внедрение планов непрерывности бизнеса.

Планы это четко определенный перечень действий и ответственных лиц, которые обеспечивают экстренное восстановление и, по возможности, нормальное функционирование компании после инцидентов/кризиса/катастрофы. Лучшая международная практика определяет три элемента для создания эффективного и действенного плана – реагирование, управление инцидентами/кризисом и восстановление деятельности.

Национальный институт стандартизации и технологий США (NIST) разработал методологию, которая определяет и описывает типовые планы обеспечения непрерывности бизнеса, уделяя особое внимание тому, что планы должны предусматривать не только предоставление технических решений, а содержать четкую организационную модель поведения в кризисных ситуациях.

VI. Тестирование планов и обучение персонала.

Тестирование и обучение являются неотъемлемыми элементами процесса непрерывности, поскольку информированность, осведомленность и практика применения планов помогают избежать определенных ошибок во время кризиса и, определенным образом, минимизировать время восстановления деятельности компании.

VII. Обновление планов.

Обновление осуществляется на регулярной основе, также в случаях изменения структуры, технических условий или требований, законодательства, выявления уязвимых мест во время тестирования.

Принимая во внимание изложенное, непрерывность бизнеса быстро становится важным и неотъемлемым элементом работы любой компании. Поэтому, стоит подчеркнуть несколько основных моментов, которые определяют направления развития этой деятельности, а также ее важные элементы.

Непрерывность бизнеса это совокупность заранее определенных действий, которые применяет компания для предотвращения и реагирования на угрозы. Указанный процесс обеспечивает способность компании непрерывно предоставлять услуги, минимизировать влияние кризисных ситуаций на свою деятельность и уменьшить возможные убытки от таких ситуаций.

Непрерывность бизнеса это комплексный механизм, работа которого требует применения не только организационных, но и технических мер. В любом случае, вопрос непрерывности является глобальным. Т.е., защита требуется компании в целом, а не отдельно выделенному элементу. Современные подходы и решения способны не только обеспечить непрерывность функционирования ИТ-ресурсов, но и обеспечить возможность, хотя иногда и ограниченную, для функционирования компании, защиты персонала и доступа к необходимым ресурсам.

Бесспорно, что при формировании процесса непрерывности одним из ключевых элементов является вопрос информационной безопасности, в частности ресурсов, используемых бизнесом (данные, аппаратные и программные комплексы, соответствующий персонал).

Определение только технических средств не обеспечит надлежащего функционирования процесса непрерывности бизнеса. Кроме того, требуется определение критических процессов бизнеса, которые и будут восстанавливаться при помощи технических средств, влияние на бизнес, проведение анализа рисков и, непосредственно, формирование самого процесса обеспечения непрерывности бизнеса.

Практически каждый этап непрерывности реализуется благодаря персоналу и сотрудникам компании. Т.е., наличие квалифицированных, подготовленных сотрудников является залогом эффективного внедрения планов непрерывности бизнеса. А обучение и информирование персонала по поводу мер и действий по восстановлению функционирования должно стать неотъемлемым элементом ежедневной работы компании.

Таким образом, обеспечение непрерывности бизнеса это важный фактор эффективного функционирования компании. Непрерывность бизнеса невозможно четко определить в системе работы компании и привязать к какой-либо определенной функции. Это комплексные меры, исполнение которых зависит от всех привлеченных сотрудников – службы безопасности, информационных технологий, информационной безопасности, логистики, персонала и др.

Необходимо помнить, что без прямого привлечения собственников процессов, планы непрерывности бизнеса будут иметь исключительно теоретический характер, и, в случае возникновения кризисной ситуации, могут только помешать, а не исполнять свою основную функцию – обеспечивать и реализовывать непрерывность бизнеса компании.

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ – О ЧЕМ ГОВОРЯТ ЭКСПЕРТЫ

Источник: http://ko.com.ua/informacionnaya_bezopasnost_o_chem_govoryat_jeksperty_117694

Компания Integrity Vision небезуспешно провела свою первую конференцию UA Security Conference, которую планирует сделать ежегодной. Ее целью, по словам организаторов, было демонстрация новейших решений в области управления информационной безопасностью, рассмотрение актуальных практик и организация площадки для обмена опытом.

Конференцию открыл генеральный директор компании Олег Половинко. Он отметил, что главной ценностью компании является ее коллектив экспертов. У Integrity Vision за всю историю ее шестилетнего присутствия на рынке нет незавершенных проектов. Второй ценностью он назвал долгосрочное партнерство. Все заказчики, которые начали работать с компанией в 2010 г., продолжают сотрудничество. Основные направления деятельности Integrity Vision – поставки оборудования, ПО и управление бизнес-процессами. С этого года в портфеле решений компании появилась информационная безопасность. Это вызвано тем, что роль ИБ в мире ИТ будет только повышаться.

Почему нас атакуют, как нас атакуют и где найти решение? На эти вопросы попытался ответить эксперт по ИБ Олег Пивовар. И своем докладе он использовал, в основном, данные из аналитических отчетов по безопасности компании Verizon Enterprise. В числе основных мотивов атак докладчик назвал деньги, промышленный шпионаж, хобби, идеологические мотивы и личную неприязнь. При этом наибольшая доля приходится на первые два. Именно они должны учитываться при построении системы ИБ в организациях.

По статистике за 1994–2014 гг. 92% инцидентов описывались девятью шаблонами атак. Столько же шаблонов применялось ив 2014–2016 гг., но уже в 95% случаев. Среди первых четырех – атаки на веб-приложения, на POS-терминалы, неправомерное использование учетных данных и физические кражи и потери. Основными направлениями атак в 2014 – 2016 гг. были веб-приложения и POS-терминалы. Наиболее часто для атак использовались уязвимости, фишинг и учетные данные. Эксперт обратил внимание аудитории на то, что 99,9% уязвимостей были использованы через год после их публикации; 23% процента пользователей открывают фишинговые электронные письма и 11% из них открывают присоединения, 50% – открывают письмо и присоединение в течение часа. Что касается учетных данных, то 60% инцидентов произошли по вине администраторов из-за неправильных настроек.

Олег Пивовар: «Наиболее часто для атак в 2014–2016 гг. использовались уязвимости, фишинг и учетные данные»

Естественным вопросом при построении системы ИБ является «с чего начать?» Своеобразное руководство под названием Top 20 Critical Security Controls опубликовало Агентство национальной безопасности США (NSA). В нем собран некий набор правил, который ранжирован по уровню критичности. Может показаться странным, но две первых позиции в нем занимают инвентаризация оборудования и ПО. И только две следующих занимают операции, непосредственно относящиеся к безопасности – это конфигурация аппаратных средств и непрерывный учет и устранение уязвимостей. Что еще важно, так это переход от реактивной защиты к проактивной, поскольку у злоумышленника всегда есть преимущество во времени.

Для того чтобы правильно организовать защиту, нужно знать все фазы выполнения атаки, их анатомию. Об этом рассказал управляющий директор компании Qualys Павел Сотников.

Павел Сотников: «Причиной осуществимости многих атак является невыполнение базовых технических процессов, как то установки обновлений, отсутствие контроля изменений, выявления инцидентов и анализа уязвимостей»

Вначале злоумышленники собирают данные о компании. Информация собирается из социальных сетей, с помощью пассивного поиска, с публичного сайта компании. На основе собранных данных они готовят инструменты, с помощью которых собираются провести атаку. Для этого могут использоваться, например, имеющиеся в свободном доступе генераторы вредоносного кода. Далее осуществляется тем или иным способом доставка вредоносного кода (фишинг, USB flash и т. д.) и используются известные или неизвестные (zero day) уязвимости. Затем устанавливается присутствие, захватывается управление, выполняются намеченные операции и скрываются следы атаки.

Далее выступающий рассмотрел несколько нашумевших атак, в том числе и BlackEnergy. Причиной осуществимости многих атак является невыполнение базовых технических процессов, как то установки обновлений, ошибки в конфигурации, отсутствие контроля изменений, выявление инцидентов и анализ уязвимостей. К этому можно приплюсовать отсутствие цикла защищенной разработки приложений, контроля защищенности поставщиков и низкую осведомленность пользователей.

Ситуативная защита сегодня не работает. Таково мнение технического директора по безопасности в регионе Центральной и Восточной Европы Анджея Клешницки (Andrzej Klesnicki) из Qualys. Многие атаки достигают цели потому, что не выполняются элементарные требования правил «гигиены» безопасности: вовремя устанавливать заплатки, изменять пароли и установки по умолчанию, удалять неиспользуемые сервисы.

Анджей Клешницки: «Не следует стараться защититься от всего – нужно оценить угрозы безопасности и определить, от каких из них необходимо защищаться»

Типичная картина, наблюдаемая в сегменте ИБ, – это много угроз и ограниченные ресурсы для построения «круговой» защиты. Возникает вопрос, как распределить имеющийся бюджет? Здесь нужно помнить, что эффективность защиты определяется ее слабейшим звеном.

Как важную меру обеспечения ИБ докладчик назвал регулярное сканирование систем для определения уязвимостей. Результаты сканирования, как правило, оформляются в виде отчетов. Однако этот метод имеет ряд недостатков. К примеру, если сканирование выполняется раз в месяц, то отчеты формируются вовремя. Но если сканирование выполняется чаще, то такой подход уже не работает. В этих случаях нужно переходить к постоянному мониторингу безопасности. Нужно постоянно оценивать все

активы, постоянно собирать информацию об инцидентах и немедленно реагировать. Для объяснения, что такое постоянный мониторинг безопасности, Анджей Клешницки привел определение NIST. Вероятно, оно будет интересно и нашим читателям: «Термины «непрерывные» и «текущие» в данном контексте означает, что безопасность и организационные риски оцениваются и анализируются на частоте, достаточной для поддержки основанных на оценке рисков решений безопасности для адекватной защиты информации организации. Сбор данных, независимо от частоты, выполняется через дискретные интервалы».

Однако не следует стараться защититься от всего. Нужно оценить угрозы безопасности и определить, от каких из них необходимо защищаться. Не нужно защищаться от того, что не является угрозой для компании. Второй момент – это ограничение по времени. Когда уязвимость становится известной, то злоумышленникам для ее изучения необходимо от 20 до 40 дней. Это то время, когда можно от нее защититься.

Сегодня известно о восьми тысячах уязвимостей. Их количество растет каждый год. Поэтому для непрерывного управления уязвимостями необходим план. Его первым пунктом должен быть постоянный сбор информации по всей инфраструктуре. Далее, нужно постоянно следить за новостями в области ИБ, иметь план установки заплаток в непредвиденных случаях, устранять другие уязвимости посредством регулярного процесса установки заплаток.

Защита от утечки данных (DLP) является важным компонентом в общей системе ИБ. О том, что предлагает в этой области компания Symantec, рассказал руководитель направления Павел Назаревич. Вначале он отметил, что компания сегодня сосредоточена только на разработках продуктов по ИБ. Два основных направления разработок – это защита инфраструктуры и безопасность контента, или непосредственно защита информации. Именно в это направление входит Symantec DLP.

Павел Назаревич: «Для надежной защиты конфиденциальной информации необходимо специализированное решение»

Каждая организация может использовать простые меры для предотвращения утечки данных, например, запретить доступ к съемным носителям, ограничить или запретить печать, шифровать данные, применять другие административные меры. Однако, по словам докладчика, это лишь частично решает проблему. Для надежной защиты данных необходимо специализированное решение.

За всеми угрозами, с которыми сталкиваются компании, стоят люди. Это могут быть добропорядочные инсайдеры, которые «не знают, что творят», злонамеренные инсайдеры или внешние охотники за конфиденциальной информацией. Согласно внутренней аналитике Symantec, 64% данных утекают случайно по вине добропорядочных инсайдеров, а 50% увольняемых сотрудников уходят с данными. В Европе средняя стоимость потерь в опрошенных компаниях составила 5,4 млн долл.

Как правило, не все данные в компании нуждаются в защите. Поэтому перед запуском DLP нужно провести классификацию данных и защищать только важные, к примеру, данные кредитных карт, медицинские и персональные данные, финансовую информацию и т. п. Начинать строить защиту необходимо с создания политик вручную или по шаблонам. Здесь можно воспользоваться аудитом либо для начала определить заведомо конфиденциальную информацию. Основными требованиями к функциональности DLP являются поиск защищаемых данных в сети и на конечных устройствах, где она не должна находиться (обнаружение), слежение за событиями и проверка пересылаемых данных (мониторинг) и защита, которая может включать блокировку, удаление, шифрование, помещение в карантин. В заключение должен проводиться анализ и предприниматься соответствующие меры по снижению рисков. Вся эта функциональность, по словам докладчика, реализована в Symantec DLP. Система защищает сеть, компьютеры, хранилища и мобильные устройства.

Для анализа контента Symantec использует инновационные технологии. Защищается не файл в целом, а конфиденциальная информация в нем. После ее идентификации, где бы она ни появилась и в каком бы формате она ни появилась, при пересылке происходит анализ, фиксируется инцидент и пересылка останавливается. Для анализа контента используются морфология и цифровые отпечатки двух видов – для структурированных данных и для неструктурированных, а также технология, которую компания называет Vector Machine Learning. Последняя позволяет обучить систему определенным правилам идентификации информации для таких сложных случаев, как исходные коды, CAD/CAM, всевозможная графическая информация.

При защите конечных точек Endpoint DLP определяет, как данные используются, куда отправляются и на каких рабочих станциях хранится конфиденциальная информация. Для этого применяются контроль сохранения, копирования и перемещения данных, контроль бизнес-процессов и сканирование рабочих станций, соответственно. Защита хранилищ (DLP for Storage) предусматривает, в частности, такие функции, как сканирование хранилищ с целью определения, какая информация, с какими правами и где размещена, и при необходимости информация перемещается. При защите сети можно контролировать интернет- и email-трафик как на уровне рабочих станций, так и на уровне сети. При этом выполняется контроль ключевых точек и магистральных каналов, контроль на шлюзах и в случае необходимости – блокировка и шифрование. Для зашифрованной информации предлагается гибкая настройка политик.

В рамках конференции состоялись две панельных дискуссии, на которых обсуждались такие темы, как ИБ глазами CIO и CISO ведущих украинских компаний и тенденции 2016–2017 в области ИБ.

Материалы конференции доступны для всех желающих по запросу на сайте мероприятия.

ЗМІСТ