

Чернышов А. В.

**ВАЖНЫЕ АСПЕКТЫ СОЗДАНИЯ ЛОКАЛЬНЫХ ДОЛГОВРЕМЕННЫХ
АРХИВНЫХ ХРАНИЛИЩ ЭЛЕКТРОННОЙ ИНФОРМАЦИИ**

Александр Викторович Чернышов

кандидат технических наук, доцент

sch-ru@yandex.ru

*Мытищинский филиал Московского государственного технического
университета имени Н. Э. Баумана*

**IMPORTANT ASPECTS OF CREATING LOCAL LONG-TERM ARCHIVAL
REPOSITORIES OF ELECTRONIC INFORMATION**

Alexandr Viktorovich Chernyshov

Mytischki Branch of Bauman Moscow State Technical University

Аннотация. В статье раскрыта проблематика создания долговременных архивных хранилищ электронной информации, показана важность создания таких хранилищ в локальном исполнении и рассмотрены важные аспекты, которые должны быть учтены при их проектировании.

Abstract. The article deals with the problem of creating long-term archival repositories of electronic information, shows the importance of creating such repositories in local variant and considers important aspects that should be taken into account when designing them.

Ключевые слова: долговременные архивы электронной информации

Keywords: long-term archives of electronic information

В последнее время в нашем обществе всё большее количество информации производится и потребляется исключительно в электронном виде. Более того, информация, накопленная обществом в «физическом» виде (преимущественно в виде записей на бумажных и иных человекочитаемых носителях), активно оцифровывается и также становится доступна в электронном виде.

Отличительной особенностью электронной информации является её «нематериальность», что приводит, с одной стороны, к простоте её копирования (размножения), к простоте и высокой скорости поиска нужных сведений, к удобству анализа и обработки, а с другой стороны, к её подверженности искажению вплоть до полной безвозвратной потери. Безвозвратная потеря информации несёт в себе серьёзные риски потерь — и экономических, и репутационных, и сиюминутных, и в долгосрочной перспективе.

В результате одной из проблем, стоящих перед современным человечеством, является проблема сохранения на длительное время накопившейся электронной информации. С этой целью создаются электронные архивы различного уровня и ёмкости [3]. Часто в качестве технической основы таких архивов предлагаются облачные технологии. Так в России одним из таких проектов является «О7» [4]. И вообще часто можно слышать мнение, что наиболее разумным способом создания архива электронной информации длительного хранения является использование облачной платформы с доступом через Интернет.

В качестве достоинств такого метода обычно приводятся следующие аргументы:

- отсутствуют серьёзные первоначальные финансовые вложения для создания необходимой инфраструктуры — можно начинать «закладку» информации на хранение практически в любой момент;
- нет необходимости заботиться о целостности хранимой информации и исправности аппаратуры архива — это обеспечит персонал «облака»;
- в процессе хранения информации необходимо вносить абонентскую плату лишь за реально сохраняемый объём информации.

В действительности перечисленные выше достоинства являются справедливыми лишь отчасти и лишь для информации, сохраняемой на краткосрочную перспективу (единицы лет).

В долгосрочной перспективе (десятки лет) могут возникнуть следующие серьёзные проблемы:

- регулярно вносимая абонентская плата за хранение накопившегося объёма информации может превысить предполагаемые затраты на содержание собственной локальной инфраструктуры хранения;

- вносить абонентскую плату необходимо регулярно, независимо от текущего финансового состояния владельца информации — даже кратковременная потеря платежеспособности может привести к безвозвратной потере хранимой информации;

- владелец облачного сервиса может потерять интерес к этому роду занятия и закрыть свой облачный сервис (возможно несколько вариантов дальнейшего развития событий, более или менее благоприятных для владельца информации, но в большинстве случаев он будет просто поставлен перед фактом);

- владелец сервиса хранения может в одностороннем порядке пересмотреть свою бизнес модель, что может привести к различным неудобствам для владельца информации, например, к повышению абонентской платы или к ограничению на скорость доступа к информации;

- в силу целого ряда причин информация может стать доступна третьим лицам без ведома её владельца;

- доступ к информации может быть прекращён/приостановлен владельцем сервиса хранения из-за каких-либо санкций.

В любом случае при использовании облачных технологий существует ещё одна проблема — передача информации на облако и обратно выполняется по сети Интернет, то есть по ненадёжным линиям связи, что может приводить к искажению информации, а наличие механизмов кеширования на различных узлах линий связи делает практически невозможным для конечного пользователя реализовать контроль её целостности в облачном хранилище. Необходимость же доступа к информации именно и только через Интернет может из-за технических сбоев приводить к полной недоступности информации, либо, как минимум, ограничивать скорость доступа.

Перечисленные выше обстоятельства показывают, что для создания долгосрочных архивов электронной информации облачные технологии являются

далеко не лучшим решением, и на первый план выходят архивы, создаваемые локально в организациях, являющихся владельцами информации. В самом деле, локальные электронные архивы имеют ряд преимуществ:

- доступ к носителям информации может быть выполнен непосредственно, что сильно упрощает контроль и целостности хранимой информации, и исправности самих носителей;
- по этой же причине доступ к любому объёму информации может быть выполнен с максимальной скоростью, допустимой для используемых носителей;
- при временных финансовых трудностях правильно сформированный электронный архив (в частности, при его создании должны быть приняты меры к дублированию информации) на некоторое время можно оставить без обслуживания — вероятность потери информации будет обусловлена лишь возможностью деградации применяемых носителей;
- вообще затраты на создание и поддержание архива могут быть конкретизированы, обоснованы и минимизированы;
- может быть реализована понятная процедура контроля доступа к информации, что минимизирует риск несанкционированного доступа.

Однако на сегодняшний день существует очень немного работ, посвящённых вопросам проектирования локальных долговременных электронных архивов, а также процессам их создания и поддержания. Поэтому существует проблема принятия обоснованных решений при проектировании таких архивов.

По мнению автора, к наиболее важным аспектам создания локальных долговременных архивных хранилищ электронной информации, научное обоснование которых требует проработки, относятся следующие.

1. Выбор типа (типов) носителей информации. Практически выбор должен быть сделан между жёсткими дисками, магнитными лентами и оптическими дисками однократной записи, причём действующий ГОСТ [1] требует использования только носителей типа WORM, что однозначно указывает на последние.

2. Обоснование количества запасных копий информации. Из-за вероятности деградации носителей информации при длительном хранении каждый носитель должен быть продублирован в нескольких запасных копиях, изменение количества которых позволяет балансировать между надёжностью хранения информации и стоимостью архива [7].

3. Выбор типа структуры, в которые объединяются носители информации в архиве. Объединение носителей в массивы типа RAID может теоретически повысить надёжность хранения информации [6], что в результате может привести к уменьшению необходимого количества запасных копий носителей (и в конечном итоге — к уменьшению затрат на создание и поддержание архива).

4. Обоснование выбора аппаратных средств, количества обслуживающего персонала и перечня регламентных работ, проводимых в архиве — все эти аспекты тесно связаны между собой и в конечном итоге определяют не только затраты на создание и содержание архива, но и предельную ёмкость архива, которая может быть достигнута без потери надёжности хранения информации, а также возможность регулярного и быстрого доступа пользователей к архивной информации.

Поясним последний пункт. Рассмотрим для примера электронный архив, создаваемый на оптических дисках однократной записи. Технически архив может быть создан как с применением стандартных приводов оптических дисков, в которые операторы загружают диски вручную, так и с применением роботизированных библиотек оптических дисков [2, 5]. Первый вариант по затратам на аппаратуру значительно дешевле второго. Однако поскольку на запись и чтение каждого оптического диска требуется вполне конкретное время, с ростом объёма архива может оказаться, что невозможно выполнять регулярный контроль целостности каждого диска с заданной периодичностью. Можно либо увеличить количество используемых приводов (и, возможно, операторов), что приведёт к удорожанию содержания архива, либо снизить частоту проверок целостности

дисков, что приведёт к снижению надёжности хранения информации. Любопытно, что целесообразность применения в этом ключе роботизированной библиотеки тоже требует обоснования.

В действительности все перечисленные выше аспекты тесно взаимосвязаны. Например, стремление выиграть в надёжности хранения информации может приводить к значительному увеличению стоимости содержания архива, а стремление минимизировать этот рост — к существенному увеличению времени доступа пользователей к информации.

Таким образом, необходима разработка метода, который позволял бы проектировать оптимальные для конкретных условий долговременные архивные хранилища электронной информации с учётом всех перечисленных аспектов.

Список литературы

1. ГОСТ Р 54989-2012 / ISO TR 18492:2005. Обеспечение долговременной сохранности электронных документов. (Вступил в силу 01.05.2013).

2. Data Archiver LB-DH8 series. – URL: <http://panasonic.net/avc/archiver/lb-dh8/> (дата обращения: 02.01.2016).

3. Егоров, Г. А. Принципы практической реализации современных архивных хранилищ данных / Г. А. Егоров, В. И. Шяудкулис, М. Финотти, М. И. Беляков // Информационные технологии и вычислительные системы. – 2013. – № 1. С. 21–32.

4. «Национальная облачная платформа» (Проект «О7») / ОАО «Ростелеком». – URL: <http://www.rostelecom.ru/projects/innovations/o7/> (дата обращения: 2016.09.29).

5. ODS-L10 Optical Disc Archive 10-Slot library. – URL: <http://www.pro.sony.eu/pro/lang/en/eu/product/archiving-storage-oda-small-library/ods-l10/overview/> (дата обращения: 2016.05.17).

6. Чернышов, А. В. Исследование свойств долговременных электронных архивных хранилищ информации на оптических дисках, организованных в структуры RAID-5 / А. В. Чернышов // Информационные технологии. – 2018. – Т. 24, – № 9. – С. 586–593. DOI: 10.17587/it.24.586-593.

7. Чернышов, А. В. К вопросу о применении оптических дисков для создания долговременных электронных архивных хранилищ информации небольших организаций / А. В. Чернышов // Информационные технологии. – 2016. – Т. 22 – № 8. – С. 635–640.