

## ЭНЕРГЕТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ПРОЦЕССА ДОЛГОВРЕМЕННОГО ХРАНЕНИЯ ДАННЫХ

Н. А. ВЕРЗУН, М. О. КОЛБАНЁВ, Е. Д. ПОЙМАНОВА

*Санкт-Петербургский государственный экономический университет,  
191023, Санкт-Петербург, Россия  
E-mail: e.d.poymanova@gmail.com*

Одна из новых задач в области информатизации — длительное сохранение больших массивов цифровых данных. Показано, что ее решение требует разработки новых технологий, отличных от тех, которые используются в операционных системах реального времени. Такие характеристики современных систем, как время записи/считывания данных или время доступа к данным, в системах долговременного хранения отступают на второй план. Одним из ключевых требований становится возможность технологии обеспечить сохранение данных в течение заданного времени при минимальных затратах энергии. Исследуется взаимозависимость энергетических и временных характеристик систем долговременного сохранения цифровых данных.

**Ключевые слова:** *данные, процесс сохранения, система хранения, энергетические ресурсы, зеленые информационные технологии, энергетический барьер, метаданные, блок данных*

Рост цифровой вселенной предъявляет новые требования к технологиям сохранения данных. Одной из новых задач в этой области информатизации стало сохранение больших информационных массивов в течение длительного времени [1, 2]. Размерность задачи может быть оценена следующим образом:

— при объеме цифровой вселенной в зета- и йотабайты размеры массивов данных, предназначенных для хранения, могут достигать не только тера-, но и петабайт;

— требуемое время хранения некоторых видов данных может измеряться не только десятками лет, как, например, при архивном хранении, но и сотнями и даже тысячами лет, если речь идет о фамильных данных или данных, потенциально влияющих на выживание будущих поколений человечества.

Можно ожидать, что в ближайшее время появятся специализированные предприятия, предлагающие соответствующие услуги [3—6].

Широко распространенные сегодня технологии сохранения данных не ориентированы на долговременное хранение [7]. Они предназначены для использования в системах реального времени [8], позволяют многократно изменять относительно небольшие объемы данных, обеспечивать их доставку пользователям по инфокоммуникационным сетям. Конкурентоспособность подобных систем хранения тем выше, чем выше плотность записи на информационный носитель и меньше время записи/считывания данных (целью разработчиков также является уменьшение времени доступа к данным). Носители для таких систем можно проектировать исходя из времени гарантированного сохранения, ограниченного годами, как правило, этого достаточно для работы систем реального времени.

Исторически долговременное (в течение тысячелетий [2]) хранение гарантировали „аналоговые технологии“, такие как глиняные таблички шумеров, египетские папирусы или русские берестяные грамоты. И в наше время информацию с ветхих бумажных носителей в архивах и библиотеках можно надежно сохранить при помощи аналоговых технологий путем, например, фотографирования и изготовления слайдов. Однако этот опыт не может быть использован при сохранении большей части цифровых данных из-за их огромного объема и

трудности или невозможности визуализации. Что же будет известно о нашей цифровой цивилизации через 5000 лет? Возможно ли сохранение цифровых данных в течение такого времени?

Одной из специфических проблем при долговременном хранении цифровых данных является сложность алгоритма кодирования цифровой информации. Так, табличку, созданную шумерами, можно „прочитать“ и сегодня, поскольку за 5000 лет органы зрения у людей не изменились. Чтение цифрового кода возможно только при доступе к эталонным моделям цифровых данных, которые позволят выделить из информационного массива метаданные и в соответствии с ними превратить цифровой код в сигналы, доступные для органов чувств человека. Необходимость сохранения эталонных моделей и метаданных совместно с данными пользователя существенно влияет на характеристики процесса сохранения. Служебная информация увеличивает объемы исходного информационного массива, подлежащего сохранению, вследствие чего ухудшаются и все технологические характеристики информационных систем (если оценивать их по отношению к полезным — пользовательским — данным).

Каждому хранимому на цифровых носителях элементу данных ставится в соответствие минимальная единица хранения (МЕХ). МЕХ — это физическая система, главное свойство которой заключается в способности оставаться в одном из нескольких (в частности, двух) устойчивых состояний [9], установленных управляющим сигналом. В магнитной памяти МЕХ — это домен, т.е. макроскопическая область, в которой вектор спонтанной однородной намагниченности определенным образом ориентирован относительно соответствующего вектора в соседних доменах. В полупроводниковой и оптической памяти — это транзисторы интегральной схемы и информационный рельеф на поверхности оптического диска соответственно. Если минимальная единица хранения имеет два устойчивых состояния, то может сохранить 1 бит данных, если четыре, то — 2 бита, если восемь, то — 3 бита и т.д.

В процессе хранения информационного носителя МЕХ может искажаться. Вероятность искажения определяется законом Аррениуса, она тем выше, чем меньше энергетический барьер, отделяющий минимальные единицы хранения друг от друга и от среды [10]. Создание сильных энергетических барьеров требует большей энергии. Эта энергия достаточно мала при записи одного бита, но растет по мере роста объемов хранения до величин, которыми уже нельзя пренебрегать.

Для сохранения больших объемов данных на длительное время нужны энергоэффективные информационные системы, обеспечивающие гарантированное сохранение данных при минимальных затратах энергии [11—13]. Это требование в полной мере соответствует и принципам построения зеленых информационных технологий [14].

В статье исследуется взаимозависимость энергетических и временных характеристик систем долговременного сохранения цифровых данных.

Введем следующие обозначения:

$u$  — энергетический барьер, отделяющий МЕХ друг от друга и внешней среды. Будем предполагать, что этот барьер одинаков у всех МЕХ и для его установки требуется энергия, равная  $u$  (Дж);

$\tau$  — время гарантированного сохранения МЕХ (лет), зависящее от  $u$ .

Значение  $\tau$  зависит от силы энергетического барьера следующим образом [15]:

$$\tau = \tau_0 \exp\left(\frac{u}{k_B T}\right), \quad (1)$$

где  $\tau_0 \sim 10^{-9}$  с для магнитных материалов,  $k_B$  — постоянная Больцмана,  $T$  — температура окружающей среды (К).

Зависимость гарантированного времени сохранения от энергетического барьера представлена на рис. 1, из которого, в частности, видно, что барьер  $u \approx 11,48k_B T$  позволит сохра-

нить данные 100 лет,  $13,78k_B T$  — 1000 лет,  $20,69k_B T$  — 1 000 000 лет. Можно также заметить, что скорость увеличения энергетического барьера растет по мере увеличения времени гарантированного сохранения.

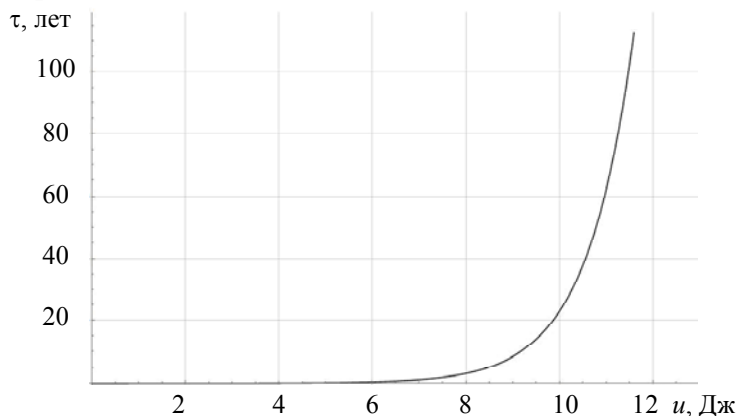


Рис. 1

Теперь оценим энергетические затраты системы сохранения за все требуемое время.

В общем случае процесс длительного хранения цифровых данных состоит из циклов перезаписи (с целью восстановления энергетических барьеров МЭХ), продолжительность которых зависит от времени гарантированного сохранения  $\tau$  [16]. Процедура перезаписи может выполняться и в ручном, и в автоматическом режиме. Количество перезаписей данных за требуемое время сохранения будет тем больше, чем меньше  $\tau$ . Уменьшить число перезаписей можно, увеличивая энергетический барьер, за счет дополнительных энергозатрат при управлении состояниями МЭХ.

Обозначим дополнительно:

$t$  — требуемое время сохранения информационного массива, устанавливаемое пользователем (заказчиком услуги);

$n$  — число циклов перезаписи за время  $t$

$$n \geq \frac{t}{\tau} \text{ или } n = \frac{t}{\tau} k, \quad (2)$$

где  $k$  — коэффициент поправки, выбираемый исходя из предположения о реальных внешних условиях хранения информационных носителей. Выполнение условия (2) обеспечит поддержание уровня энергетического барьера до того момента времени, когда истекает время гарантированного сохранения МЭХ;

$E$  — энергия, затрачиваемая за время  $t$ :

$$E = nu = nk_B T \ln \left( \frac{\tau}{\tau_0} \right). \quad (3)$$

На рис. 2 представлена зависимость энергии от требуемого времени хранения с учетом реалистичных значений  $n$ .

Видно, что существует такое значение энергетического барьера, при котором общие энергетические затраты за время директивного сохранения минимальны.

Технологии сохранения никогда „не работают“ с одним битом данных или с целым информационным массивом. Они обеспечивают разбиение исходного информационного массива, подлежащего сохранению, на отдельные блоки данных, причем разбиение не связано со смыслом информации. Например, в операционных системах реального времени требуется, чтобы магнитный информационный носитель на аппаратном уровне был разделен на отдельные сектора. Каждый сектор помимо данных из исходного информационного массива должен содержать дополнительные метаданные, которые позволяют обнаружить его на носителе,

произвести считывание и коррекцию данных и другие операции [17—19]. Далее на программном уровне операционная система формирует более крупные блоки данных и связанные с ними дополнительные метаданные. Иерархия блоков данных в общем случае может иметь несколько уровней.

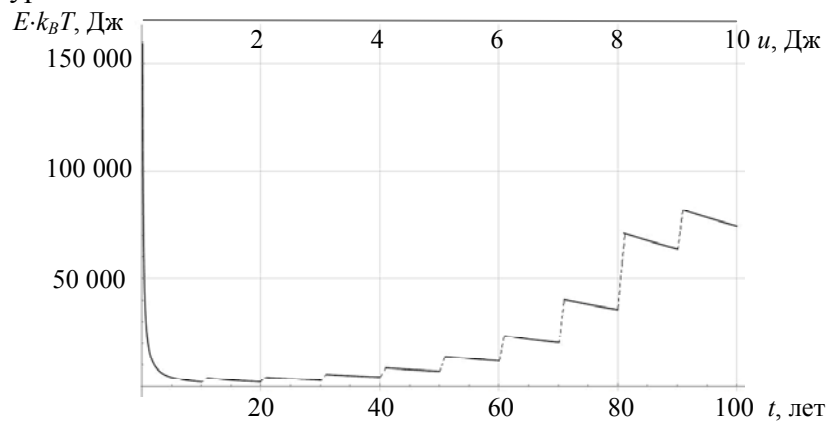


Рис. 2

Для упрощения расчетов далее будем предполагать, что выделение блоков данных осуществляется на одном уровне. Обозначим:

$V$  — число битов данных на носителе;

$S$  — число битов данных в одном блоке;

$S_M$  — количество метаданных в одном блоке;

$S_{\Pi}$  — количество „полезных“ данных в одном блоке:

$$S = S_M + S_{\Pi};$$

$P$  — число блоков на носителе:

$$P = V/S;$$

$E_T$  — энергия, затрачиваемая за время гарантированного хранения, с учетом технологии записи;

$E_{\Pi}$  — энергия, потраченная на запись „полезных“ данных;

$E_M$  — энергия, потраченная на запись метаданных:

$$E_T = E_{\Pi} + E_M.$$

Очевидно, что энергия, затрачиваемая на сохранение исходного информационного массива, зависит от размера блока данных. При увеличении последнего будет уменьшаться количество метаданных, приходящихся на каждый бит „полезной“ информации или, что то же самое, уменьшаться энергия, затрачиваемая на запись метаданных.

Вернемся к процедуре записи. Каждый блок состоит из „полезных“ данных и метаданных, содержащих служебную информацию. Количество метаданных, приходящихся на один блок, мало зависит от количества „полезных“ данных и может считаться константой. Таким образом, можно предположить, что чем больше размер записываемых блоков на носителе, тем меньше доля метаданных на нем, соответственно для сохранения одного „полезного“ бита потребуется меньше энергии.

Размер записываемых блоков данных определяется требованиями операционной системы. Очевидно, что с увеличением размера блока данных скорость обработки информации операционной системой будет падать. Однако для системы долговременного хранения скорость обработки не является критичной, поскольку к данным не требуется обращение в режиме реального времени.

С учетом емкости сохраняемого информационного массива общая энергия  $E$ , затраченная на запись  $V$  бит, будет равна:

$$E = k_B T \ln \left( V \frac{t}{\tau_0} \right). \quad (5)$$

Энергия, затрачиваемая на запись „полезных данных“, на магнитный диск размером  $V$  бит:

$$E_{\Pi} = k_B T \ln \left( S_{\Pi} P \frac{t}{\tau_0} \right) = k_B T \ln \left( (S - S_M) \frac{V}{S} \frac{t}{\tau_0} \right), \quad (6)$$

на запись метаданных:

$$E_M = k_B T \ln \left( S_M P \frac{t}{\tau_0} \right) = k_B T \ln \left( S_M \frac{V}{S} \frac{t}{\tau_0} \right). \quad (7)$$

На рис. 3 показана зависимость энергии от размера записываемого блока данных. Как и следовало ожидать, энергия, затрачиваемая на сохранение исходного информационного массива, существенно сокращается при увеличении размера блока данных и уменьшении объема метаданных.

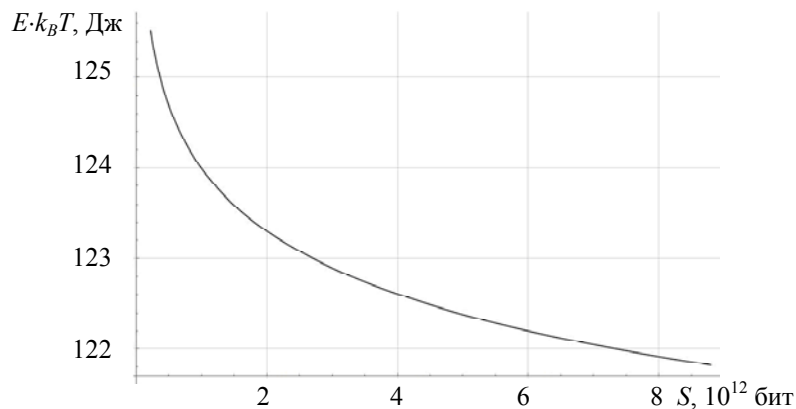


Рис. 3

Результаты исследования показывают, что для долговременного хранения больших массивов данных необходимы новые технологии и системы. Главным требованием к ним является не способность работать в масштабе реального времени, а возможность малых энергозатрат при длительном сохранении. Как показывают проведенные расчеты, при организации информационного процесса долговременного хранения данных имеется возможность настроить характеристики технологии таким образом, чтобы уменьшить общее потребление энергетических ресурсов.

#### СПИСОК ЛИТЕРАТУРЫ

1. Советов Б. Я., Колбанёв М. О., Татарникова Т. М. Диалектика информационных процессов и технологий // Информация и космос. 2014. № 3. С. 96—104.
2. Советов Б. Я., Колбанёв М. О., Татарникова Т. М. Информационные процессы и технологии. СПб: ГУАП, 2014.
3. Колбанёв М. О., Микадзе С. Ю., Татарникова Т. М. Модель информационного взаимодействия для предприятий сервиса // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 10—14.
4. Воробьёв А. И., Колбанёв М. О., Татарникова Т. М. Оценка вероятностно-временных характеристик процесса предоставления информационно-справочных услуг // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 15—18.
5. Кориунов И. Л. Состояние и концепция развития информационных технологий в сфере сервиса // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 7—10.
6. Головкин Ю. Б., Ярцев Р. А., Газетдинова С. Г. Моделирование процессов инструментальной подготовки сервисного обслуживания на основе экспертных оценок // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 27—32.

7. Татарникова Т. М., Пойманова Е. Д. Технологии долговременного хранения данных // Сб. науч. тр. по матер. Междунар. науч.-практ. конф. „Наука и образование в XXI веке“. 2013. С. 136—138.
8. Богатырев В. А., Богатырев А. В., Богатырев С. В. Оценка надежности выполнения кластерами запросов реального времени // Изв. вузов. Приборостроение. 2014. Т. 57, № 4. С. 46—48.
9. Колбанёв М. О., Татарникова Т. М. Информационный объем базовых информационных процессов // Информационно-управляющие системы. 2014. № 4. С. 42—47.
10. Kish L.B., Granqvist C.G. Does Information Have Mass? // Proc. IEEE. 2013. Vol. 101, N 9. P. 1895—1899.
11. Колбанёв М. О., Пойманова Е. Д., Татарникова Т. М. Физические ресурсы информационного процесса сохранения данных // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 38—42.
12. Колбанёв М. О., Татарникова Т. М. Физические ресурсы информационных процессов и технологий // Науч.-техн. вестн. информационных технологий, механики и оптики. 2014. Т. 14, № 6. С. 113—122.
13. Верзун Н. А., Колбанёв М. О., Омелян А. В. Об энергетической эффективности сетей пакетной передачи данных // Изв. вузов. Приборостроение. 2014. Т. 57, № 9. С. 42—46.
14. Воробьев А. И., Колбанев А. М., Колбанев М. О. Экологическая безопасность информационных технологий // Геополитика и безопасность. 2015. Т. 32, № 4. С. 90—99.
15. Мейлихов Е. З., Фарзетдинова Р. М. Максимальная плотность магнитной записи и распределение переключающих полей // Физика твердого тела. 2014. Т. 56, вып. 12. С. 2326—2334.
16. Колбанев М. О., Татарникова Т. М., Малков К. О. Подход к организации адаптивного согласующего центра корпоративной сети // Информационно-управляющие системы. 2008. № 3. С. 28—31.
17. Bogatyrev V. A. Protocols for dynamic distribution of requests through a bus with variable logic ring for reception authority transfer // Automatic Control and Computer Sciences. 1999. Vol. 33, N 1. P. 57—63.
18. Богатырев В. А., Богатырев А. В. Оптимизация резервированного распределения запросов в кластерных системах реального времени // Информационные технологии. 2015. Т. 21, № 7. С. 495—502.
19. Татарникова Т. М., Пойманова Е. Д. Роль метаданных при длительном хранении информации // Сб. науч. тр. по матер. Междунар. науч.-практ. конф. „Современное общество, образование и наука“. 2015. С. 141—142.

**Сведения об авторах**

- Наталья Аркадьевна Верзун** — канд. техн. наук, доцент; СПбГЭУ, кафедра информационных систем и технологий; E-mail: dina\_25@hotmail.ru
- Михаил Олегович Колбанёв** — д-р техн. наук, профессор; СПбГЭУ, кафедра информационных систем и технологий; E-mail: mokolbanev@mail.ru
- Екатерина Дмитриевна Пойманова** — аспирант; СПбГЭУ, кафедра прикладных информационных технологий; E-mail: e.d.poymanova@gmail.com

Рекомендована кафедрой  
информационных систем и технологий  
СПбГЭУ

Поступила в редакцию  
29.06.16 г.

**Ссылка для цитирования:** Верзун Н. А., Колбанёв М. О., Пойманова Е. Д. Энергетические характеристики процесса долговременного хранения данных // Изв. вузов. Приборостроение. 2017. Т. 60, № 2. С. 158—164.

**ENERGY PERFORMANCE OF LONG-TERM DATA STORAGE PROCESS**

**N. A. Verzun, M. O. Kolbanev, E. D. Poimanova**

*St. Petersburg State University of Economics, 191023, St. Petersburg, Russia*  
E-mail: e.d.poymanova@gmail.com

The problem of long-term storage of large amounts of digital data is considered. Solutioning to the problem is shown to call for development of new technologies different from those used in real-time operation systems. Characteristics of modern systems, such as time of data recording/reading or time of data access in long-term storage systems recede into the background, while one of the key requirements becomes the ability of technology to provide persistent storage for the specified time at minimum cost of energy. Interdependence of energy and time characteristics of systems long-term storage of digital data are studied.

**Keywords:** data, storage system, energy sources, green information technology, energy barrier, metadata, data unit

**Data on authors**

- Natalya A. Verzun** — PhD, Associate Professor; St. Petersburg State University of Economics, Department of Information Systems and Technologies;  
E-mail: dina\_25@hotmail.ru
- Mikhail O. Kolbanev** — Dr. Sci., Professor; St. Petersburg State University of Economics, Department of Information Systems and Technologies;  
E-mail: mokolbanev@mail.ru
- Ekaterina D. Poimanova** — Post-Graduate Student; St. Petersburg State University of Economics, Department of Information Systems and Technologies;  
E-mail: e.d.poymanova@gmail.com

**For citation:** Verzun N. A., Kolbanev M. O., Poimanova E. D. Energy performance of long-term data storage process // Journal of Instrument Engineering. 2017. Vol. 60, N 2. P. 158—164 (in Russian).

DOI: 10.17586/0021-3454-2017-60-2-158-164