

## РЕФЕРАТ

**Актуальність теми.** Незважаючи на наявність деякої кількості робіт, що розглядають тег-орієнтовані файлові сховища та системи, досі не було запропоновано методу побудови такого сховища з використанням прямого доступу до накопичувача, без проміжної ієрархічної файлової системи. Використання прямого доступу до накопичувача дозволяє виключити залежності від певного типу проміжної системи та виконати додаткову оптимізацію роботи файлового сховища. Метод побудови сховища з прямим доступом до накопичувача також дозволяє створити файлову систему на його основі. Запропонований метод є внеском у вирішення науково-технічної проблеми, актуальної для широкого спектру областей інформатики, адже файлове сховище чи система є невід'ємною компонентою будь-якої комп'ютерної системи.

**Об'єктом дослідження** є процеси організації асоціативно індексованого зберігання інформації.

**Предметом дослідження** є побудова тег-орієнтованого файлового сховища на накопичувачі з використанням прямого доступу, а також управління вже побудованим таким сховищем.

**Мета роботи:** створення методу, що дає змогу виконати побудову тег-орієнтованого файлового сховища на накопичувачі з використанням прямого доступу, без необхідності використання проміжної ієрархічної файлової системи, а також управління таким сховищем в подальшому.

**Наукова новизна** полягає в наступному:

1. Створено метод побудови тег-орієнтованого файлового сховища на накопичувачі з використанням прямого доступу.

2. Розроблено спосіб пакування часток файлу в незаповнений блок накопичувача.

**Практична цінність** отриманих в роботі результатів полягає в тому, що запропонований метод побудови тег-орієнтованого файлового сховища на накопичувачі з використанням прямого доступу, без необхідності використання проміжної ієрархічної файлової системи, яке може бути використане для збереження та управління наборами файлів. Теоретичні та практичні результати роботи можуть бути використані в організаціях, що потребують збереження, обробки та пошуку у великих об'ємах складно структурованих файлів.

**Апробація роботи.** Основні положення і результати роботи були представлені та обговорювались на науковій конференції магістрантів та аспірантів “Прикладна математика та комп'ютинг” ПМК-2018 (Київ, 21-23 березня 2018 р.) та на науковій конференції магістрантів та аспірантів “Прикладна математика та комп'ютинг” ПМК-2018-2 (Київ, 14-16 листопада 2018 р.)

**Структура та обсяг роботи.** Магістерська дисертація складається з вступу, чотирьох розділів та висновків.

*У вступі* подано загальну характеристику роботи, зроблено оцінку сучасного стану проблеми, обґрунтовано актуальність напрямку досліджень, сформульовано мету і задачі досліджень, показано наукову новизну отриманих результатів і практичну цінність роботи, наведено відомості про апробацію результатів.

*У першому розділі* розглянуто проблематику та сценарії використання файлових сховищ та систем, існуючі методи їх вирішення, а також проведений аналіз, який дає змогу визначити основні переваги та недоліки цих методів. Розглянуті характеристики, за якими можна виконати

порівняння таких методів. Виявлено не вирішені, або недостатньо повно вирішені, проблеми.

*У другому розділі* наведено опис розробленого методу побудови тег-орієнтованого файлового сховища, формату накопичувача, таблиць індексації.

*У третьому розділі* описані структура і алгоритми програмної реалізації методу у вигляді модуля (бібліотеки) а також еталонної реалізації програмної системи, що використовує його. Наведені рекомендації щодо використання програмної системи та інтеграції модуля в інші програмні системи.

*У четвертому розділі* проводиться аналіз ефективності запропонованого методу та його еталонної реалізації в рамках виявлених у першому розділі проблем.

*У висновках* представлені результати проведеної роботи, дані рекомендації щодо використання методу.

Робота представлена на 90 аркушах, містить посилання на список використаних літературних джерел.

**Ключові слова:** тег, файлове сховище, накопичувач.

## ABSTRACT

**Relevance of the topic.** Notwithstanding the existence of a few works on tag-based file storage, as of today no method of building such storage on a mass storage device, employing direct access to said device, without intermediary hierarchic file system, was proposed. Employment of the direct access to the mass storage device allows to eliminate dependencies of certain types of intermediary file system and to perform tag-based storage-specific optimizations. Method for the building of the tag-based storage directly on the mass storage device also allows the creation of a file system based on said storage in the course of further research. Proposed method contributes to the solution of the scientific and technical problem relevant for the wide plethora of fields in informatics, as file storage is an essential component of every computer system.

**The object of research** are the processes of organization of the associatively indexed information storage.

**The subject of research** is the building of tag-based file storage directly on mass storage device, employing direct access to said device, and also management of such storage previously built.

**The purpose of the work:** creation of a method that would allow to build tag-based file storage directly on mass storage device, employing direct access to said device, without intermediary hierarchic file system, and also further management of such storage.

**The scientific novelty** is as follows:

1. A method for building of tag-based file storage directly on mass storage device, employing direct access to said device, was created.

2. The method for packing parts of a file into a partially filled block of the mass storage device was created.

**The practical value** of the results obtained in the work is that a method of building of tag-based file storage directly on mass storage device, employing direct access to said device, without intermediary hierarchic file system, that can be used for storage and management of file sets, is proposed. Theoretical and practical results of the work may be used by the organizations that require storage, processing and searching in vast volume of complexly structured files.

**Approbation of the work.** The main provisions and results of the work were presented and discussed at the scientific conference of masters and postgraduates “Applied Mathematics and Computing”, PMK-2018 (Kyiv, March 21-23, 2018) and at the scientific conference of masters and postgraduates “Applied Mathematics and Computing”, PMK-2018-2 (Kyiv, November 14-16, 2018).

**Structure and scope of the work.** The master’s dissertation consists of an introduction, four sections and conclusions.

*The introduction* gives a general description of the work, assesses the current state of the problem, substantiates the relevance of the research direction, formulates the purpose and objectives of the research, shows the scientific novelty of the results obtained and the practical value of the work, provides information on the approbation of the results.

*The first section* examines the issues and scenarios of the file system and storage usage, existing methods of solution of the issues described, and also provides the analysis that allows to determine the main advantages and disadvantages of said methods. Yet unsolved, or not solved in complete entirety, problems are revealed.

*The second section* describes the proposed method for building of the tag-based storage, mass storage device format and the indexation tables.

*The third section* describes the structure and algorithms of the implementation of the proposed method in the form of a loadable software module (library) and the reference implementation of a software system employing the said module. Recommendations regarding the usage of the software system and the integration of the module into other software systems, are provided.

*The fourth section* performs the analysis of the proposed method's (and its reference implementation's) efficiency within the confines of the problems described in the first section.

*In the conclusions* the results of the work are presented, along with the recommendations regarding the usage of the method.

The work spans 90 sheets and contains references to the literary sources used.

**Keywords:** tag, file storage, mass storage device.

## РЕФЕРАТ

**Актуальность темы.** Несмотря на наличие некоторого количества работ, рассматривающих тег-ориентированные файловые хранилища и системы, до сих пор не было предложено метода построения такого хранилища с использованием прямого доступа к накопителю, без промежуточной иерархической файловой системы. Использование прямого доступа к накопителю позволяет исключить зависимости от определённого типа промежуточной системы и выполнить дополнительную оптимизацию работы файлового хранилища. Метод построения хранилища с прямым доступом к накопителю также позволяет создать файловую систему на его основе. Предложенный метод является вкладом в решение научно-технической проблемы, актуальной для широкого спектра областей информатики, ведь файловое хранилище или система является неотъемлемой компонентой любой компьютерной системы.

**Объектом исследования** являются процессы организации ассоциативно индексируемого хранения информации.

**Предметом исследования** является построение тег-ориентированного файлового хранилища на накопителе с использованием прямого доступа, а также управление уже построенным таким хранилищем.

**Цель работы:** создание метода, позволяющего выполнить построение тег-ориентированного файлового хранилища на накопителе с использованием прямого доступа, без необходимости использования промежуточной иерархической файловой системы, а также управление таким хранилищем в дальнейшем.

**Научная новизна** заключается в следующем:

1. Создан метод построения тег-ориентированного файлового хранилища на накопителе с использованием прямого доступа.

2. Разработано способ упаковывания частиц файла в незаполненный блок накопителя.

**Практическая ценность** полученных в работе результатов заключается в том, что предложен метод построения тег-ориентированного файлового хранилища на накопителе с использованием прямого доступа, без необходимости использования промежуточной иерархической файловой системы, которое может быть использовано для хранения и управления наборами файлов. Теоретические и практические результаты работы могут быть использованы в организациях, требующих хранения, обработки и поиска в больших объёмах сложно структурированных файлов.

**Апробация работы.** Основные положения и результаты работы были представлены и обсуждались на научной конференции магистрантов и аспирантов «Прикладная математика и компьютеринг» ПМК-2018 (Киев, 21-23 марта 2018 г.) и на научной конференции магистрантов и аспирантов «Прикладная математика и компьютеринг» ПМК-2018-2 (Киев, 14-16 ноября 2018 г.).

**Структура и объём работы.** Магистерская диссертация состоит из введения, четырёх глав и выводов.

*В введении* представлена общая характеристика работы, сделана оценка современного состояния проблемы, обоснована актуальность направления исследований, сформулирована цель и задачи исследований, показана научная новизна полученных результатов и практическая ценность работы, приведены сведения об апробации результатов.

*В первой главе* рассмотрены проблематика и сценарии использования файловых хранилищ и систем, существующие методы их решения, а также проведён анализ, позволяющий определить основные преимущества и



недостатки этих методов. Рассмотрены характеристики, по которым возможно выполнить сравнение таких методов. Определены не решённые, либо недостаточно полно решённые, проблемы.

*Во второй главе* приведено описание разработанного метода построения тег-ориентированного файлового хранилища, формата накопителя, таблиц индексации.

*В третьей главе* описаны структура и алгоритмы программной реализации метода в виде модуля (библиотеки), а также эталонной реализации программной системы, использующей его. Приведены рекомендации по использованию программной системы и интеграции модуля в иные программные системы.

*В четвёртой главе* проведён анализ эффективности предложенного метода и его эталонной реализации в рамках определённых в первой главе проблем.

*В выводах* представлены результаты проведённой работы, даны рекомендации по применению метода.

Работа представлена на 90 листах, содержит ссылки на список использованных литературных источников.

**Ключевые слова:** тег, файловое хранилище, накопитель.