

С. П. Белов

МАКЕТИРОВАНИЕ ПРОИЗВОДИТЕЛЬНОСТИ ФАЙЛОВОЙ СИСТЕМЫ NTFS ПРИ НАГРУЗКЕ ТИПА «ЗАПИСЬ ДАННЫХ НА НАКОПИТЕЛЬ SSD» И РАЗНЫХ РАЗМЕРАХ КЛАСТЕРА NTFS

Введение

Изобретение твердотельных накопителей SSD позволило преодолеть некоторые недостатки магнитных носителей HDD. Избавление от движущихся элементов внутри конструкции, казалось бы, избавило SSD от каких-либо задержек записи информации, особенно с применением технологий TRIM, и NCQ. Но было не совсем ясно, влияет ли размер кластера файловой системы на скорость записи данных.

Существующие работы [1 – 4], посвященные обзору производительности NTFS, имеют уклон в теоретическую часть и описывают зависимости большим формульным набором. Также неизвестны

алгоритмы тестирования этих зависимостей. Авторы статей не выходят на связь по электронной почте. Поэтому цель данной статьи — показать зависимости файловой системы NTFS с точки зрения чисел, а не формул, и при этом описать четкие алгоритмы работы анализирующего ПО и ТТХ оборудования.

Макетирование

Для макетирования было создано ПО, предназначенное для многократного копирования файла-эталона на целевой носитель и измерения затрачиваемого времени. Причины создания именно своего ПО были следующие:

использование устаревшей среды программирования Microsoft Visual Basic v.6.0 как создателя дистрибутива, не требовательного к ресурсам [5];

исключение обновления элементов интерфейса ПО, как замедляющего фактора работы ПО [6];

реализация собственных алгоритмов работы с данными; программы вида SpeedTest показывают скорость работы объекта, но это число остается труднопроверяемым и принимаемым на веру. Во время разработки же своего ПО каждый новый нюанс работы ОС или носителя вносил существенные коррективы в полученные выводы, так как резко менялась интерпретация результатов. Поэтому открытые алгоритмы работы с данными необходимы.

Отладка ПО проводилась записью с RAM-диска на RAM-диск и заключалась в следующем:

приближении к идентичности времени записи в рамках порций;

приближении к идентичности времени записи в рамках количества проходов: на каждый проход тратится одинаковое время;

приближении к идентичности времени записи в рамках расчетов: вариаций количества файлов в порции, порций в проходе и количества проходов.

Отладка ПО выявила проблемы и результаты:

операционная система (интерфейс) вносит коррективы в копирование файлов в виде дополнительных задержек, если использовать ее встроенный функционал «копировать-вставить». Копирование в Total Commander той же информации происходит быстрее. В созданном ПО — еще быстрее (интерфейс Total Commander также вносит задержки). Значит идея самостоятельно написанного ПО подтверждена, и одним из его свойств должно быть полное «зависание» ПО на всем интервале времени работы с файлами;

использование WinAPI-функции CopyFile в составе библиотеки kernel32.dll оказалось верным решением, приведшим к успешности отладки. Использование класса FileSystemObject в составе библиотеки scrrun.dll (Microsoft Scripting Runtime) не оправдалось: хаотичные задержки при обработке порций файлов;

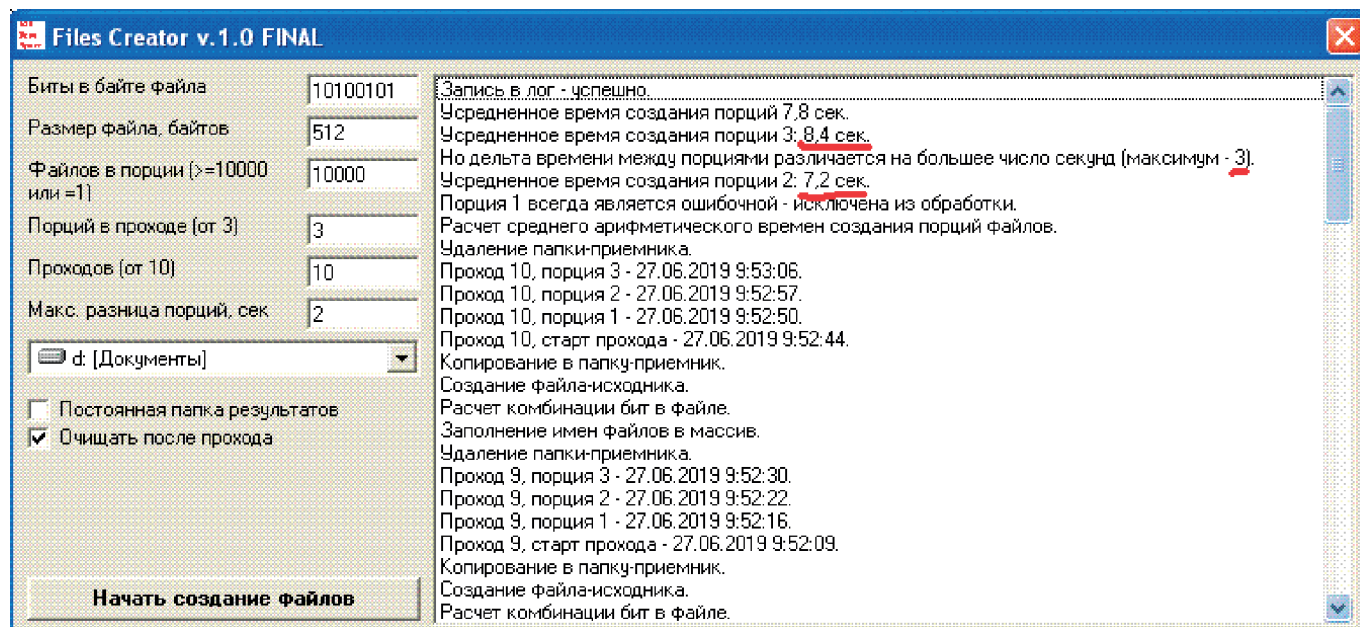


Рис. 1. Результат работы ПО на HDD с уровнем фрагментации 15% и заполнением 56% (данные ПО «Defraggler»). Имеется сильное расхождение во времени записи порций файлов с одинаковым номером, что сказывается на усредненном времени между порциями с разными номерами

при записи в RAM зафиксирована одинаковая скорость записи в кластер и 1 байта, и полного размера кластера — для всего разнообразия кластеров; присутствуют флуктуации между результатами одного теста; запись той или иной порции файлов различается по времени при большом количестве подходов, что сказывается на среднем результате погрешностью до 5% / тест. Это связано с тем, что Windows не является системой реального времени и что при высоко нагруженной системе часы могут отставать на секунды, а затем резко восстанавливаться до нормального значения. Возможно, десятикратное увеличение порции файлов еще более сгладит данную погрешность, однако каждый тест в этом случае будет многодневным;

файловая система NTFS является замедлителем записи файлов на носитель при увеличении количества файлов в папке, но только после превышения предела, заключенного в промежутке [200; 300] тысяч файлов, увеличивающего итоговое время записи от 23,7%; незначительное замедление записи файлов происходит при пересечении отметки 3,2 млн. файлов в разделе, увеличивающее итоговое время записи от 2,7%;

для получения более точных результатов использовалась программа Process Lasso Pro — выделить ПО отдельно 4-е ядро и запретить использование этого ядра другим программам; однако даже это не спасало от увеличения времени записи файлов при использовании сторонним ПО других ядер (от 6,5%); для форматирования раздела использовалась программа Partition Wizard Home Edition: встроенный механизм форматирования ОС содержит ошибку (форматировал раздел 2048 Б как 1024 Б);

SSD для страховки выровнена и разбита на разделы — именно в Windows 7 SP1 до использования в целевой ОС (реализация Advanced Format без повреждения носителя при работе в ОС старше Windows 7 SP1) [7].

Блок-схема алгоритмов работы ПО представлена на рис. 2. Исходный код представлен как ссылка на ПО [8].



Рис. 2. Блок-схема алгоритмов работы ПО Тестер файловых систем v.1.0 FINAL (название изменено в последний момент)

Технические характеристики оборудования, используемого при работе с ПО Тестер файловых систем v.1.0 FINAL (название изменено в последний момент).

Тестовый объект SSD:

60ГБ Silicon Power V30 SSD SATA3, файловая система NTFS;

запуск ПО – с виртуального RAM-диска (из оперативной памяти) для исключения влияния системного носителя на тестовый носитель (даже копирование файла размером 1 байт с флешки приводит к существенному искажению измерений). Программа виртуального диска – WinRamTech (бывшая QSoft) RamDrive Enterprise. Приоритет приложения – высокий, насколько позволяет операционная система.

Тестовый стенд №1 (тестовый объект – RAM).

Комплектующие: AMD Sempron 3300+ 2ГГц, Gigabyte GA-K8VM800M SATA1, 1ГБ DDR1 Elixir 400МГц, 120ГБ Seagate ST3120213A HDD IDE;

операционная система: Microsoft Windows 7 Максимальная x32;

настройки системы: отключены службы дефрагментации, Windows Search, брандмауэра, контроля учетных записей, центра обеспечения безопасности, архивации, обновления; отключены опции сжатия данных и индексирования содержимого файлов на носителях; отключены файлы гибернации и подкачки; отключена заставка и гашение экрана;

настройки BIOS: режим SATA - IDE, технология NCQ – отсутствует.

Тестовый стенд №2 (тестовый объект – RAM).

Комплектующие: Intel Duo E4500 2.2ГГц, ASRock ConRoe133-DVI-H SATA2, 2ГБ DDR2 (разных номина-

Сергей Павлович Белов – ведущий инженер-программист АО «Государственный научно-исследовательский институт приборостроения».

лов, частот и производителей), 40ГБ Samsung SP0411N HDD IDE;

операционная система: Microsoft Windows XP Профессиональная x32;

настройки системы: отключены службы брандмауэра, центра обеспечения безопасности, архивации, обновления; отключены опции сжатия данных и индексирования содержимого файлов на носителях; отключены файлы гибернации и подкачки; отключена заставка и гашение экрана;

настройки BIOS: режим SATA - IDE, технология NCQ – отсутствует.

Тестовый стенд №3 (тестовый объект – RAM, SSD).

Комплектующие: AMD A10-5700 3,39 ГГц, Gigabyte F2A75M-D3H SATA3, 32ГБ DDR3 Corsair 1600 МГц, 500 ГБ Plextor PX-512M5Pro SSD SATA3;

операционная система: Microsoft Windows XP Профессиональная x64;

настройки системы: отключены службы брандмауэра, центра обеспечения безопасности, архивации, обновления; отключены опции сжатия данных и индексирования содержимого файлов на носителях; отключены файлы гибернации и подкачки; отключена заставка и гашение экрана;

настройки BIOS: режим SATA - IDE, технология NCQ – отключена.

* При каждом форматировании диска Windows включает индексирование диска автоматически -- требуется отключение вручную.

Дальнейшая отладка ПО была направлена на измерение времен записи уже на SSD. Каждый тест запускался 3 раза, вычислялось среднее значение. 16 проходов, 20 порций файлов по 10000 штук – 3,2 млн. файлов за тест. Результаты тестирования не смогли быть использованы по назначению, но были внесены уточнения по минимальным системным требованиям ПО:

тестовый стенд №№ 1 и 2: индикатор загруженности SSD не горит ярко постоянно; постоянность ярко свечения имеет периодичность, сменяясь неярким свечением, порой угасая полностью; скорость записи файлов на SSD и RAM-диск одинаково некорректная: не хватает частоты процессора 2,2 ГГц – данные тестовые стенды использовать невозможно;

дальнейшие успешные тесты на тестовом стенде № 3 показали, что данное одноядерное ПО требует процессор с большой тактовой частотой, как самый слабый и нагруженный элемент тестового стенда; минимальная частота определена экспериментально

как 3,4 ГГц: на тестовом стенде № 3 скорость записи порции одних и тех же файлов занимала одинаковое время, индикатор загруженности SSD горел ярко постоянно большую часть времени, тускло постоянно меньшую часть времени. В работе автора Е.И. Нижник также отражено существенное влияние процессора на корректность результатов тестирования файловой системы NTFS [9, 10];

оперативной памяти требуется не менее 400МБ частотой не менее 533МГц.

Дальнейшие тесты приводили уже к анализу именно тестового носителя SSD на тестовом стенде № 3 с настройками: 10000 файлов в порции, 20 порций, 16 проходов, трехкратное выполнение. Была составлена сводная табл. 1, отражающая среднее время записи порции файлов.

Анализ данных табл. 1 позволяет сделать следующие выводы:

скорость обработки данных высока, если сравнивать запись целиком одного кластера и двойного;

Таблица 1*

УСРЕДНЕННОЕ ВРЕМЯ ЗАПИСИ ФАЙЛОВ В РАЗДЕЛ С РАЗНЫМИ РАЗМЕРАМИ КЛАСТЕРА (СЕКУНД НА 10000 ФАЙЛОВ)

Размер кластера	Минимум (1 байт)	Минимальный излишек (1 кластер + 1 байт)	Файлы размером одного кластера - в другие кластеры								1 большой файл 20ГБ**
			512	1024	2048	4096	8192	16384	32768	65536	
512	3,271	3,299	3,272	3,327	3,414	3,465	3,504	3,64	4,004	4,867	119,864
1024	3,197	3,3	3,23	3,211	3,359	3,358	3,38	3,54	3,868	4,643	121,564
2048	3,283	3,48	3,336	3,356	3,302	3,525	3,544	3,663	4,041	4,788	120,615
4096	3,399	3,454	3,404	3,409	3,444	3,441	3,463	3,635	4,031	4,826	120,805
8192	3,162	3,295	3,186	3,181	3,215	3,226	3,206	3,373	3,703	4,159	***
16384	3,01	3,226	3,024	3,02	3,014	3,087	3,073	3,147	3,459	3,733	***
32768	2,888	3,324	2,904	2,906	2,913	2,943	2,954	2,961	3,26	3,458	***
65536	2,878	3,84	2,876	2,882	2,908	2,911	2,944	2,957	3,175	3,37	***

* Подробные данные см. в табл. 2.

** 3 файла в порции из-за малого размера носителя (при таком объеме файла и времени его записи - межфайловые задержки ничтожны).

*** Носитель вышел из строя: превышен предел записи для ячеек.

разница в десятые-сотые доли секунды; при этом запись одного кластера занимает единицы секунд, т. е. имеется проблема задержек в файловой системе, а не в контроллере диска;

скорость обработки огромного файла 20ГБ в разных кластерах одинакова с учетом погрешности (влияние файловой системы в данный момент минимально), т. е. для больших файлов размер кластера ФС не имеет никакого значения, а термин «большой» можно численно уточнить в будущем, как меньше 20 ГБ;

нужно ввести такое понятие, как скорость записи кластерной пустоты; если сравнивать запись одного байта в кластер и целиком кластера в кластер — время записи растет от 0,001с для 512 Б до 0,492 с для 64 КБ; при этом удельная скорость записи с / Б не имеет четкой зависимости от размера кластера и лежит в диапазоне $[1,956 \cdot 10^{-6}; 13,685 \cdot 10^{-6}]$ с / Б. Таким образом, оптимальность выбора кластера для ФС является математической задачей определения количества кластерной пустоты в существующих файлах, количества файлов и их размеры, расчета времени записи пустоты по усредненной скорости $7,8205 \cdot 10^{-6}$ с / Б, т. е. определение «чем больше кластер ФС, тем быстрее», является неверным. Так же как и рекомендация Microsoft использовать размер файловой системы 4096 Б (указывается при форматировании с параметрами по умолчанию);

разница между записью 1 Б и 512 Б в кластер 512 Б — ничтожна и составляет 0,001 с; можно было бы считать критерием оптимальности «одинаковость времени записи независимость от кластерной пустоты», однако это всего лишь совпадение (погрешность измерений, см. табл. 2);

постепенное незначительное повышение времени, затрачиваемое на запись файлов больше 1 Б, так же говорит о проблеме задержек именно в файловой системе, а не в контроллере диска;

разница между минимальным и максимальным временем записи кластера целиком составляет 1,992 с на 10000 файлов. Несмотря на внушительное различие в 69,23%, стоит задаться вопросом о рациональности применения понятия «оптимальный размер кластера» при работе с файловой системой NTFS; потери времени и труда, затраченные на вычисление оптимального размера кластера, могут превысить выгоду в скорости записи информации.

Заключение

Произведено макетирование скорости записи файлов различного размера в раздел с различными размерами кластера. Полученные значения времен показывают:

проблему задержки записи информации именно по вине NTFS, а не контроллера диска;

неприменимость понятия «оптимальный размер кластера» для файлов большого объема из-за одинаковых времен записи в разные кластеры, с учетом погрешности измерений;

применимость понятия «оптимальный размер кластера» для файлов небольшого объема — только математическими расчетами, включающими количество файлов, их размер, количество пустых байтов в кластерах. На практике это означает установку данных в раздел с произвольным размером кластера, анализ раздела специфическим ПО, сохранение файлов в другом месте, переформатирование раздела под нужный размер кластера, возвращение файлов в раздел;

ввиду того, что требуется фактический расчет количества пустых байтов в кластерах для получения оптимального размера кластера, — файлы должны существовать в разделе физически или быть известными заранее. Следовательно, рекомендация Microsoft по использованию «стандартного размер кластера» как более быстрого [11, 12] не является

НОВЫЕ РАЗРАБОТКИ

Таблица 2

ПОДРОБНАЯ ТАБЛ. 1: ВСЕ ПОЛУЧЕННЫЕ ИЗМЕРЕНИЯ С НОСИТЕЛЯ SSD, ПОКА ОН НЕ ВЫШЕЛ ИЗ СТРОЯ

Размер кластера	Минимум	Минимальный излишек	Файлы размером одного кластера - в другие кластеры								1 большой файл*
	1 байт	1 кластер + 1 байт	512	1024	2048	4096	8192	16384	32768	65536	20ГБ
512	3,271 (3,299, 3,266, 3,247)	3,299 (3,253, 3,309, 3,335)	3,254 (3,260, 3,250, 3,253)	3,327 (3,322, 3,299, 3,362)	3,427 (3,467, 3,391, 3,424)	3,443 (3,437, 3,434, 3,457)	3,488 (3,487, 3,543, 3,434)	3,628 (3,628, 3,599, 3,658)	4,013 (4,036, 4,007, 3,997)	4,862 (4,819, 4,859, 4,908)	119,417 (118,188, 120,813, 119,250)
512 (случайные повторы)	-	-	3,289 (3,296, 3,289, 3,283)	-	3,401 (3,385, 3,414, 3,405)	3,487 (3,497, 3,493, 3,470)	3,519 (3,493, 3,553, 3,510)	3,651 (3,658, 3,651, 3,645)	3,995 (4,049, 3,980, 3,957)	4,872	120,312
1024	3,184 (3,181, 3,191, 3,181)	3,280 (3,257, 3,266, 3,316)	3,224 (3,234, 3,227, 3,211)	3,192 (3,220, 3,181, 3,174)	3,352 (3,359, 3,316, 3,382)	3,366 (3,345, 3,349, 3,405)	3,316 (3,247, 3,332, 3,368)	3,533 (3,510, 3,539, 3,549)	3,941 (3,918, 3,918, 3,988)	4,681 (4,724, 4,628, 4,691)	121,875 (121,438, 121,250, 122,938)
1024 (случайные повторы)	3,209 (3,25 3,194, 3,184)	3,32 (3,319, 3,332, 3,309)	3,235 (3,214, 3,243, 3,247)	3,229 (3,24, 3,237, 3,211)	3,365 (3,365, 3,368, 3,362)	3,35 (3,332, 3,359, 3,359)	3,442 (3,461, 3,421, 3,444)	3,546 (3,523, 3,569)	3,794 (3,855, 3,734)	4,604 (4,694, 4,569, 4,549)	121,252 (121,250, 120,750, 121,431, 121,247, 121, 582)
2048	3,283 (3,303, 3,280, 3,266)	3,480 (3,451, 3,493, 3,497)	3,336 (3,352, 3,342, 3,313)	3,356 (3,306, 3,398, 3,365)	3,302 (3,283, 3,316, 3,306)	3,525 (3,516, 3,543, 3,516)	3,544 (3,579, 3,533, 3,520)	3,663 (3,625, 3,658, 3,707)	4,041 (4,086, 4,023, 4,013)	4,788 (4,832, 4,783, 4,750)	120,615 (120,621, 120,722, 120,502)
4096	3,405 (3,411, 3,398, 3,405)	3,443 (3,451, 3,461, 3,418)	3,404 (3,408, 3,395, 3,408)	3,409 (3,411, 3,408, 3,408)	3,444 (3,385, 3,495, 3,451)	3,434 (3,437, 3,428, 3,438)	3,463 (3,49, 3,454, 3,444)	3,635 (3,635, 3,641, 3,628)	4,031 (4,03, 4,03, 4,033)	4,826 (4,878, 4,78, 4,819)	119,867 (118,467, 119,393, 121,742)
4096 (случайные повторы)	3,392 (3,398, 3,395, 3,382)	3,466 (3,451, 3,477, 3,47)	-	-	-	3,448 (3,431, 3,451, 3,461)	-	-	-	-	121,743
8192	3,162 (3,164, 3,164, 3,158)	3,295 (3,299, 3,286, 3,299)	3,186 (3,178, 3,168, 3,211)	3,181 (3,187, 3,178, 3,178)	3,215 (3,178, 3,274, 3,194)	3,226 (3,243, 3,257, 3,178)	3,206 (3,224, 3,191, 3,204)	3,373 (3,365, 3,401, 3,352)	3,703 (3,707, 3,678, 3,724)	4,159 (4,211, 4,158, 4,109)	-**
16384	3,010 (3,016, 2,993, 3,02)	3,235 (3,234, 3,237)	3,024 (3,033, 3,007, 3,033)	3,020 (2,997, 3,046, 3,016)	3,014 (3,046, 3,023, 2,974)	3,087 (3,122, 3,142, 3,043)	3,073 (3,122, 3,036, 3,062)	3,147 (3,164, 3,148, 3,128)	3,459 (3,477, 3,467, 3,434)	3,733 (3,734, 3,714, 3,750)	-**
16384 (случайные повторы)	-	3,218 (3,224, 3,243, 3,227, 3,201, 3,197)	-	-	-	-	-	-	-	-	-
32768	2,888 (2,895, 2,878, 2,891)	3,324 (3,309, 3,355, 3,309)	2,904 (2,895, 2,895, 2,921)	2,906 (2,895, 2,908, 2,914)	2,913 (2,914, 2,938, 2,888)	2,943 (2,934, 2,967, 2,928)	2,954 (2,960, 2,973, 2,930)	2,961 (2,987, 2,954, 2,941)	3,26 (3,299, 3,247, 3,234)	3,458 (3,477, 3,451, 3,447)	-**
65536	2,878 (2,868, 2,882, 2,885)	3,84 (3,842, 3,842, 3,836)	2,876 (2,878, 2,878, 2,872)	2,882 (2,878, 2,895, 2,872)	2,908 (2,882, 2,928, 2,914)	2,911 (2,934, 2,918, 2,882)	2,944 (2,954, 2,954, 2,924)	2,957 (2,967, 2,954, 2,951)	3,175 (3,197, 3,168, 3,161)	3,370 (3,477, 3,312, 3,322)	-**

*3 файла в порции из-за малого размера носителя (при таком объеме файла и времени его записи – межфайловые задержки ничтожны).

**Носитель вышел из строя: превышен предел записи для ячеек.

ванной, так как оптимальный размер кластера не зависит от объема диска.

ЛИТЕРАТУРА

1. *Семахина С.Ю.* Файловые системы FAT и NTFS операционной системы Windows . — Уфа: Академия педагогических идей «Новация», № 12, 2007.

2. *Болотов А.С., Болотова Т.П.* Файловая система NTFS: обзор версий, производительность. — М.: ФГБОУ ВПО ГУУ, Вестник Университета, № 7, 2012.

3. *Нижник Е.И., Тормасов А.Г.* Математическое моделирование производительности файловой системы NTFS при нагрузке типа «запись дисковых данных». — Новосибирск: НГУ, Вестник новосибирского государственного университета. Сер. Информационные технологии. — Т. №5, №2, 2007.

4. *Нижник Е.И.* Математическая модель нагрузки файловой системы NTFS при активном поиске дисковых данных. — М.: МФТИ, Моделирование процессов обработки информации, 2007.

5. *Сергеев В.* Visual Basic 6.0. Наиболее полное руководство для профессиональной работы в среде Visual Basic 6.0. — СПб.: ВВХ-Санкт-Петербург, 1998. — С. 8.

6. Google Developers Training team. Advanced Android Development - Practicals / USA: Google LLC, 2018 — p. 422. [Электронный ресурс] URL: [https://google-developer-training.github.io/android-developer-advanced-course-practicals.pdf](https://google-developer-training.github.io/android-developer-advanced-course-practicals/android-developer-advanced-course-practicals.pdf).

7. Компания HGST. Поддержка Advanced Format — оптимальная работа с дисками высокой емкости / М.: Hitachi Global Storage Technologies, 2013. [Электронный ресурс] URL: <https://habr.com/ru/company/hgst/blog/190564/>.

8. *Белов С.П.* Тестер файловых систем v.1.0 FINAL / М.: личный сайт. [Электронный ресурс] URL: http://www.bad-good.ru/programs.html#tester_fs.

9. *Нижник Е.И.* Роль нагрузки процессора и фрагментации диска в моделировании производительности файловой системы NTFS. -- Системы управления и информационные технологии, № 2.1 (28), 2007.

10. *Нижник Е.И.* Особенности исследования нагрузки процессора и фрагментации диска в моделировании производительности файловой системы NTFS. — Воронеж: Информационные технологии моделирования и управления, № 4 (38), 2007.

11. UA-Hosting Company. Какие факторы влияют на производительность систем хранения и как? — Латвия, Вентспилс: SIA «UA-Hosting», 2016. [Электронный ресурс] URL: <https://habr.com/ru/company/ua-hosting/blog/282469/>.

12. Microsoft. Размер кластера по умолчанию для файловых систем FAT, NTFS и exFAT / США, Вашингтон: Редмонт, официальный сайт Microsoft, 2018. [Электронный ресурс] URL: <https://support.microsoft.com/ru-ru/help/140365/default-cluster-size-for-ntfs-fat-and-exfat>.



Все
о «Петербургском
журнале электроники»
читайте на сайте

www.elstandart.spb.ru