

АЛГОРИТМ МНОГОПАРАМЕТРИЧЕСКОГО РАНЖИРОВАНИЯ ЖЕСТКИХ ДИСКОВ ПО РИСКУ ОТКАЗА

И.Н. Насыров (ФГАОУ ВО КФУ), И.И. Насыров (АО «Автоградбанк»),
Р.И. Насыров (ООО «Мединдекс»)

Рассмотрено изменение от времени эксплуатации значений SMART-параметров, характеризующих надежность жестких магнитных дисков в ПК. Выбраны параметры, являющиеся критическими в том смысле, что если их значения увеличиваются, то возрастает вероятность выхода из строя накопителей информации на жестких дисках. Научной задачей исследования является установление в отказавших жестких дисках приоритетности указанных параметров с целью разработки алгоритма оценки надежности накопителей информации различных производителей. В ходе исследования были проанализированы накопители разных производителей, эксплуатирующиеся в крупнейшем коммерческом data-центре Backblaze. В результате анализа выявлена приоритетность критических параметров. Учтено, что само наличие значений рассмотренных параметров зависит от производителя жестких дисков. Предложен алгоритм ранжирования накопителей информации по степени надежности с применением выявленной в результате исследования приоритетности параметров.

Ключевые слова: информация, накопитель, жесткий диск, надежность, ранжирование, алгоритм.

Введение

Для обеспечения сохранности информации необходимо своевременно и в полном объеме копировать данные с ненадежного накопителя на новый и надежный. С этой целью обычно используют технологию SMART (self-monitoring, analysis and reporting technology — технологию самоконтроля, анализа и отчетности) для внутренней оценки состояния жесткого диска компьютера, а также как способ предсказания возможного выхода его из строя.

В работе рассмотрено изменение от времени эксплуатации значений SMART-параметров, характеризующих надежность жестких магнитных дисков в компьютерах. Выбраны параметры, являющиеся критическими в том смысле, что если их значения увеличиваются, то возрастает вероятность выхода из строя накопителей информации на жестких дисках. Научной задачей исследования является установление в отказавших жестких дисках приоритетности указанных параметров с целью разработки алгоритма оценки надежности накопителей информации различных производителей.

В ходе исследования были проанализированы накопители торговых марок HGST (Hitachi Global Storage Technologies), Hitachi (позднее HGST), Samsung, ST (Seagate), Toshiba, WDC (Western Digital), эксплуатировавшихся в крупнейшем коммерческом data-центре Backblaze. В результате анализа выявлена следующая приоритетность критических параметров (в порядке убывания): 5 Reallocated sectors count (число переназначенных секторов), 196 Reallocation event count (число успешных и неуспешных попыток пере-

назначения), 197 Current pending sector count (число секторов, являющихся кандидатами на замену), 1 Read error rate (частота ошибок при чтении данных с диска), 9 Power-on hours (число часов, проведенных во включенном состоянии), 7 Seek error rate (частота ошибок при позиционировании блока магнитных головок), 3 Spin-up time (время раскрутки пакета дисков из состояния покоя до рабочей скорости), 10 Spin-up retry count (число повторных попыток раскрутки дисков до рабочей скорости в случае, если первая попытка была неудачной).

Учтено, что само наличие значений рассмотренных параметров зависит от производителя жестких дисков. Предложен алгоритм ранжирования накопителей информации по степени надежности с применением выявленной в результате исследования приоритетности параметров.

Методы

Для анализа зависимости значений параметров от времени эксплуатации у вышедших из строя накопителей информации на жестких магнитных дисках были изучены SMART-данные, приведенные на сайте компании Backblaze [1]. Рассмотрено 45 параметров SMART 92530 накопителей 93 моделей шести торговых марок HGST, Hitachi, Samsung, ST, Toshiba, WDC за период с 10 апреля 2013 г. по 31 декабря 2016 г. Обнаружено, что на конец исследуемого периода продолжали нормально работать 79,58% накопителей, были досрочно сняты с эксплуатации 14,74%, вышли из строя 5,68%.

Всего доступны сведения о смысловом значении более 80 параметров SMART из 256, однако боль-

шинство из них производителями не используются. Поэтому специалисты Backblaze записывали в 2013–2014 гг. только 40 из них, а начиная с 2015 г. — 45 с номерами 1–5, 7–13, 15, 22, 183, 184, 187–201, 220, 222–226, 240–242, 250–252, 254, 255 (в 2015 г. добавили 22, 220, 222, 224, 226).

Как оказалось, имеется весьма ограниченный набор параметров, которые можно применять для диагностики и оценки состояния накопителей любых производителей [2]. В первую очередь для этого было предложено использовать число переназначенных секторов. Но не как отдельный единичный параметр для оценки надежности, а как совокупность данных: текущее значение, средняя скорость накопления с момента ввода диска в эксплуатацию, мгновенная скорость изменения количества переназначенных секторов с момента последнего измерения. Аналогичное сочетание средней и мгновенной скорости изменения значений параметров используют специалисты Backblaze [3].

Доказательство приоритетности числа переназначенных секторов при оценке состояния жесткого диска представлено в [4], где приведены результаты исследования по 100 тыс. накопителей в серверах по всему миру, выполненного компанией Google.

Дополнительные параметры можно подобрать двумя разными способами: по относительным [5] и абсолютным значениям [6].

Подобный подход позволяет:

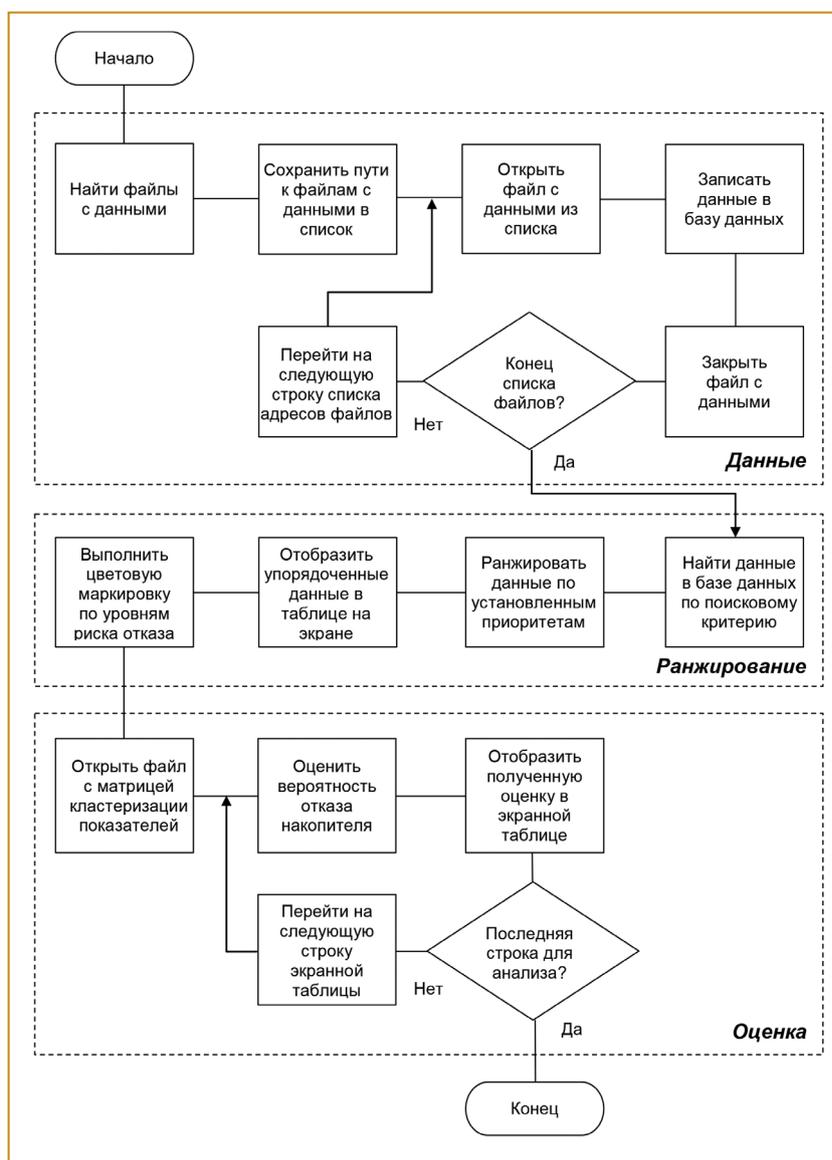
- 1) отслеживать накопители, в которых текущее значение близко к предельному уровню;
- 2) держать на контроле накопители, которые медленно, но стабильно разрушаются;
- 3) принимать экстренные меры для накопителей, в которых разовый скачок количества переназначенных секторов вызывает опасения.

Результаты и обсуждение

Первоначально был рассмотрен базовый алгоритм стандартных операций при отказе накопителей информации и обоснована необходимость разработки на его основе алгоритма оценки и прогнозирования надежности накопителей в системах управления с указанием включаемой функциональности. Первая версия алгоритма предназначалась для однопараметрической оценки и прогнозирования надежности накопителей информации. В последующей версии рассматривались уже несколько параметров.

Для многопараметрического алгоритма (рисунок) предлагается применить опыт специалистов Backblaze, которые используют для получения SMART-данных программу Smartmontools. Затем они добавляют несколько элементов, таких как дата измерения, модель накопителя, серийный номер, емкость, признак отказа и создают строку в дневном журнале регистрации для каждого диска. Эти файлы журналов можно скачать с сайта компании. Данные дисков, которые помечены в какой-то из дней как отказавшие, в журналы со следующего дня и далее не входят. Иногда диск удаляется из эксплуатации, даже несмотря на то, что он не вышел из строя, например, когда обновляется модуль хранения данных и диски объемом 1 ТБ заменяются дисками объемом 4 ТБ. В этом случае диск в 1 ТБ не помечается как сбойный, но его SMART-данные больше не регистрируются.

С одной стороны, добавление новых повторяющихся элементов, таких как модель накопителя, се-



Укрупненный алгоритм многопараметрической оценки надежности накопителей информации с ранжированием по степени вероятности отказа

рийный номер, емкость, кажутся излишними, с другой — они не позволяют данным смешиваться или теряться. Поэтому предлагаем воспользоваться опытом Backblaze и формировать файлы регистрации состояния дисков с указанными дополнениями.

Еще одна проблема связана с разрывом по времени между отказом диска и его заменой. Дело в том, что признак отказа (единица — если отказал, ноль — в остальных случаях) ставится вручную после замены диска в модуле. А до этого он может стоять там несколько дней. При этом данные с него регистрироваться не будут (он ведь не работает), хотя признак будет показывать, что он все еще работоспособный. Вследствие этого приходится также вручную заполнять образовавшиеся пустые места последними актуальными данными или изменять дату выхода из строя. Такой порядок необходим для обеспечения возможности обработки данных стандартными программами типа Excel.

Алгоритм оценки надежности накопителей информации начинается с поиска файлов с данными и сохранения путей к ним в список. Это делается для исключения повторных обращений к ним. Затем открывается первый файл данных из списка. Эти данные записываются в базу данных. Файл данных закрывается. Проверяется, достигнут ли конец списка данных. Если нет, то осуществляется переход к следующей строке списка с адресами файлов данных. Если да, то переходят к анализу и ранжированию данных.

При анализе предлагается учесть приоритет для пользователей как данных о надежности, так и вида отказов жестких дисков. Наиболее приоритетной является группа параметров, связанных с состоянием пространства памяти: 5 Reallocated sectors count (число переназначенных секторов), 196 Reallocation event count (число успешных и неуспешных попыток переназначения), 197 Current pending sector count (число секторов, являющихся кандидатами на замену), 1 Read error rate (частота ошибок при чтении данных с диска) [7]. Эти параметры имеются в наличии всегда для всех накопителей всех типов почти всех производителей (196 отсутствует у жестких дисков Samsung и Seagate (торговая марка ST)).

Вторым по приоритетности обязательно надо применить стандартное определение надежности как времени наработки до отказа, чему соответствует всегда имеющийся в наличии параметр 9 Power-on hours (число часов, проведенных во включенном состоянии) [8].

Третьей по приоритетности является группа параметров, связанных с позиционированием головки записи/считывания: 7 Seek error rate (частота ошибок при позиционировании блока магнитных головок), 2 Throughput performance (общая производительность диска), 8 Seek time performance (средняя производительность операции позиционирования магнитными головками) [9]. Ошибки позиционирования головки являются основной причиной переназначения секторов и последующего выхода из строя дисков ST. Од-

нако параметр 7 имеет нулевые значения или отсутствует у дисков Samsung и Toshiba (возможно из-за малой статистики), а параметры 2 и 8 отсутствуют у дисков Samsung, ST, WDC. Поэтому в дальнейшем рассматривается только параметр 7 как самостоятельный показатель надежности.

Четвертой по приоритетности является группа параметров, связанных с механикой приводов вращения жестких дисков: 10 Spin-up retry count (число повторных попыток раскрутки дисков до рабочей скорости в случае, если первая попытка была неудачной), 3 Spin-up time (время раскрутки пакета дисков из состояния покоя до рабочей скорости), 4 Start/stop count (полное число циклов запуск/остановка шпинделя), 12 Device power cycle count (число полных циклов включения/выключения диска), 192 Power-off retract count (число циклов выключений или аварийных отказов (включений/выключений питания накопителя)), 193 Load/unload cycle (число циклов перемещения блока магнитных головок в парковочную зону/в рабочее положение) [10]. Параметры 4, 12, 192, 193 являются накапливаемыми и зависят в основном от числа включений, в связи с чем при больших значениях те изменения, которые обусловлены другими причинами, становятся на общем фоне малозаметными. При анализе параметра 3 для оценки вероятности отказа важным является его отклонение в сторону увеличения значений. Накапливаемый параметр 10, который зависит только от состояния механики привода дисков, полностью подходит для оценки вероятности выхода их из строя по этой причине. Поэтому для указанной группы далее используются параметры 3 и 10.

Из базы данных считывается информация о параметрах на последнюю дату. Полученные данные сортируются по убыванию. Затем данные помечаются графически цветом для наглядности визуализации состояния накопителя. Для цветовой градации используются границы уровней интенсивности отказов с учетом обнаруженного факта переназначения секторов целыми дорожками (кратно 8).

Как было отмечено выше, наблюдались случаи, когда отказавшие накопители несколько дней не заменялись. Это жесткие диски моделей HGST HMS5C4040BLE640 (1 ед.), HGSTHMS5C4040ALE640 (5 ед.), ST4000DM000 (3 ед.), все емкостью 4 ТБ. При этом значения параметров не фиксировались, хотя признаки отказа показывали, что накопители работоспособны. Еще выявлены случаи, когда в процессе работы у дисков моделей WDC WD10EADS емкостью 1 ТБ (2 ед.) и ST4000DM000 емкостью 4 ТБ (52 ед.) в один из дней никаких данных не было указано. Также был выявлен случай сбоя фиксации параметров за два дня до полного выхода из строя жесткого диска модели ST4000DM000 емкостью 4 ТБ, при котором параметр 1 был пуст, а остальные имели произвольные значения. У диска модели WDC WD1600AAJS емкостью 0,16 ТБ вообще с самого начала отсутствовали данные по параметру 194. Поэтому нужно организо-

вать проверку для таких ситуаций и подготовить соответствующие операции по устранению последствий.

Следовательно, исходя из смысла и наличия значений параметров при ранжировании накопителей по степени вероятности отказа, нужно использовать сортировку сначала по параметрам 5, 196, 197 и 1, затем по 9, далее по 7, 3 и 10.

Произвести оценку вероятности выхода накопителя из строя можно также с помощью нейронных сетей, а результаты для удобства использования сгруппировать в матричном виде. В настоящее время в разных странах мира данное направление активно изучается различными группами ученых.

Открывается файл с предварительно выполненной для ускорения работы алгоритма матричной кластеризацией показателей. Производится оценка вероятности выхода из строя первого после ранжирования накопителя. Эта оценка по накопителю выводится в общую таблицу. Проверяется, достигнута ли последняя строка таблицы. Если нет, то осуществляется переход к следующей строке. Если да, то работа алгоритма завершается.

Выводы

Показано, что на базе выявленных приоритетов параметров надежности, значения которых имеются в наличии для жестких магнитных дисков всех производителей, можно выработать алгоритм оценки надежности накопителей информации с ранжированием по степени вероятности отказа.

Аргументом в обоснование применения указанных параметров является то, что для характеристики состояния пространства памяти точно такие же параметры практически с тем же самым смыслом применяются в твердотельных накопителях информации. Естественно параметры позиционирования головок записи/считывания или механизмов вращения дисков в твердотельных накопителях не используются, хотя нумерация остается прежней. Поэтому в перспективе, при переходе крупных data-центров к хранению информации в накопителях указанного типа, никаких дополнительных изменений кроме исключения параметров 3, 7 и 10 в разработанный алгоритм вносить не придется.

Принципиальное отличие рассмотренного подхода от того, что предлагают разработчики аналогов, заключается в том, что оператор видит на экране не только одну цифру вероятности выхода из строя каждого накопителя, а сразу все накопители, рассортированные по группам риска отказа. Сначала по параметру наивысшего риска, но если встречаются накопители с одинаковыми значениями этого параметра, то вторично отсортированные по следующему по приоритетности риска параметру и т. д. Когда

у наиболее приоритетного по риску отказа параметра значения становятся нулевыми, то таким же образом сортируются накопители по следующему по приоритетности параметру. В конечном итоге достаточно одного взгляда на экран, чтобы только по расположению, цвету и величине соответствующих столбцов с накопителями оценить общую картину состояния сразу всего пространства памяти в data-центре.

Заключение

Таким образом, изложенный в статье алгоритм обладает научной новизной и позволяет задачу как индивидуальной, так и общей оценки надежности накопителей информации решать с применением выявленной приоритетности параметров. Получено свидетельство о государственной регистрации программы для ЭВМ № 2018615650 от 14.05.2018 на программу многопараметрического нейросетевого ранжирования накопителей информации по степени надежности.

Список литературы

1. Backblaze. Hard Drive Data and Stats. URL: <https://www.backblaze.com>
2. Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Study of Failure Hazard Degree in Large Data Centers // Helix. 2019. Vol. 9. Is. 5. PP. 5345-5349.
3. Klein A. Hard Drive Reliability Stats for Q1 2015. URL: <https://www.backblaze.com>
4. Pinheiro E., Weber W.-D., Barroso L.A. Failure Trends in a Large Disk Drive Population // Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST'07). San Jose, California, USA, 2007. PP. 17-28.
5. Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Data mining for information storage reliability assessment by relative values // Engineering and Technology (UAE). 2018. Vol. 7. Is. 4.7 Special issue 7. PP. 204-208.
6. Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Parameters selection for information storage reliability assessment and prediction by absolute values // Advanced Research in Dynamical and Control Systems. 2018. Vol. 10. Is. 2 Special issue. PP. 2248-2254.
7. Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Reallocated sectors count parameter for analysing hard disk drive reliability // Computational and Theoretical Nanoscience. 2019. Vol. 16. Is. 12. PP. 5298-5302.
8. Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Dependence of reallocated sectors count on HDD power-on time // Engineering and Technology (UAE). 2018. Vol. 7. Is. 4.7 Special issue 7. PP. 200-203.
9. Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Positioning errors indication by Seek error rate and other HDD parameters // Advanced Research in Dynamical and Control Systems. 2019. Vol. 11. Is. 8 Special issue. PP. 1797-1805.
10. Nasyrov I.N., Nasyrov I.I., Nasyrov R.I., Khairullin B.A. Spin retry count relation with other HDD parameters // Journal of Computational and Theoretical Nanoscience. 2019. Vol. 16. Is. 12. PP. 5303-5306.

Насыров Искандар Наилович — докт. экон. наук, профессор, ФГАОУ ВО «Казанский (Приволжский) федеральный университет», **Насыров Ильдар Искандарович** — канд. техн. наук, руководитель отдела технических систем и телекоммуникаций, АО «Автоградбанк», **Насыров Рустам Искандарович** — руководитель проектов, ООО «Мединдекс».