

УДК 004.451.51

**Процесс восстановления данных в RAID-1 на основе SPDK****Сичкар Г. К., Кононова Ю. А., Васенина А. И.***Санкт-Петербургский государственный университет**199034, г. Санкт-Петербург, Университетская наб., 7–9**georgy.sichkar@yandex.ru*

Реализован процесс восстановления данных для имплементации RAID-1 внутри фреймворка SPDK, реализующего логику блочных устройств в пространстве пользователя. Реализация выполняет учет поврежденных областей RAID-1 и восстанавливает их, не прибегая к полному копированию.

**Ключевые слова:** хранение данных; RAID; дисковые сбои; восстановление данных; SPDK.

Одними из важнейших характеристик СХД являются надежность и отказоустойчивость. Иными словами, способность системы хранения в случае частичной поломки иметь возможность восстановить потерянные данные и продолжить функционировать в штатном режиме.

Примером технологии с подобной функциональностью является RAID-1 [1] — объединение нескольких блочных устройств, каждое из которых хранит копию одних и тех же данных. За счет избыточности RAID-1 позволяет восстановить информацию при замене поврежденного базового устройства.

На данный момент большинство RAID-массивов реализуются внутри пространства ядра, в том числе RAID-1. Однако возможен альтернативный подход, при котором вся основная логика СХД перемещается в пространство пользователя. Такой подход используется в SPDK [2] — фреймворке с открытым исходным кодом для работы с блочными устройствами, разработанном Intel Corporation. К особенностям SPDK можно отнести опрос оборудования на предмет завершения запросов ввода-вывода вместо использования прерываний, а также перемещение всех необходимых драйверов блочных устройств в пространство пользователя. Более того, опираясь на отчеты о производительности компании Intel, можно зафиксировать увеличение скорости работы СХД [3].

Фреймворк содержит внутри себя реализацию массива RAID-1. Однако эта реализация неполная и содержит ряд критических недочетов, влияющих на уровень предоставляемой отказоустойчивости:

- отсутствие функциональности частичного восстановления поврежденных данных, например, в ситуации, когда блочное устройство подключено по сети и временно теряет соединение;
- запуск нового процесса восстановления данных в отдельном потоке. Это может влиять на производительность всей системы, так как за время нахождения RAID-1 в активном режиме процесс восстановления может повторяться многократно, а значит, и новый поток под эти цели будет выделяться многократно.

Предложенная альтернативная реализация [4] процесса восстановления решает часть вышеперечисленных проблем.

- Реализована функциональность, отвечающая за контроль актуальности хранимых данных, позволяющая точно восстанавливать поврежденные

области. Для контроля используется матрица ребилда — структура данных, представляющая собой массив из атомарных переменных. Каждая переменная отвечает за актуальность набора областей в RAID-массиве. Для изменения статуса области используется операция CAS (compare and set), что позволяет корректно обновлять матрицу при многопоточной обработке запросов.

- Для запуска нового процесса восстановления данных применяется механизм поллинга, предоставляемый фреймворком SPDK. Этот механизм работает в одном отдельном потоке на протяжении всей работы RAID-1 и в случае появления поврежденных данных запускает процесс восстановления.
- При каждом запуске создается снимок текущего состояния RAID-массива. Восстановление данных происходит согласно снимку, в момент времени восстанавливается только фиксированный набор областей, что позволяет снизить нагрузку на систему. По окончании восстановления набора областей состояние матрицы ребилда обновляется путем выполнения операции CAS с соответствующей атомарной переменной.

#### *Литература*

1. Patterson D. A., Gibson G., Katz R. H. A case for redundant arrays of inexpensive disks (RAID) // Proceedings of the 1988 ACM SIGMOD international conference on Management of data, 1988. P. 109–116.
2. Accelerating your NVMe drives with SPDK. URL: <https://spdk.io/doc/about.html> (дата обращения: 2024.07.05).
3. Latecki K., Chachulski J. SPDK NVMe BDEV Performance Report Release 23.09. 2023. URL: [https://ci.spdk.io/download/performance-reports/SPDK\\_nvme\\_bdev\\_perf\\_report\\_2309.pdf](https://ci.spdk.io/download/performance-reports/SPDK_nvme_bdev_perf_report_2309.pdf) (дата обращения: 2024.07.05).
4. Rebuild process. URL: <https://github.com/sumr-proj/spdk/pull/11> (дата обращения: 2024.07.05).