

# Быстрый программный RAID

Обеспечивает минимальную просадку производительности массива при отказе дисков и существенно сокращает время ребилда

## Высокая скорость расчета контрольных сумм

В основе наших решений лежит программный RAID, способный производить расчет контрольных сумм быстрее, чем любые аналогичные решения в индустрии. RAID-массив читает и записывает блоки четности с рекордной скоростью (около 25 ГБ/с на 1 ядро процессора), что позволяет сохранить высокую производительность массива даже в режиме восстановления.

При отказе диска на последовательных запросах снижение скорости чтения с массива не составит более 10%, что значительно меньше, чем у любой другой системы хранения данных.

## Сокращение времени ребилда RAID-массива

Ребилд (реконструкция, восстановление) RAID-массива после отказа накопителя — это потенциально опасный промежуток времени, который может принести администратору СХД массу неудобств.

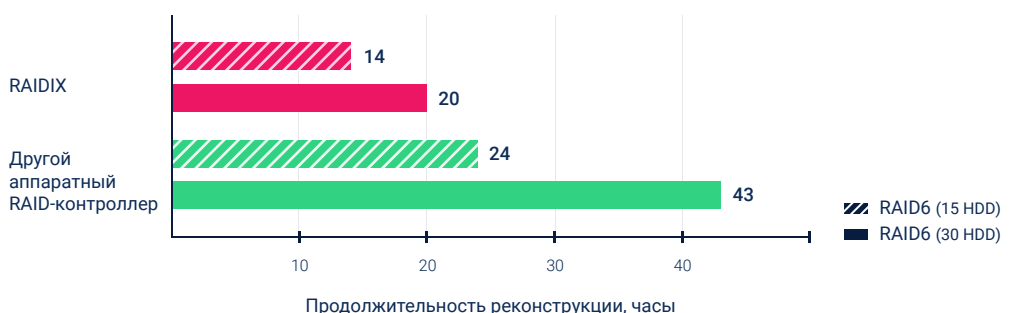
Во-первых, восстановление данных на новый диск обычно ощутимо снижает общую производительность массива. Во-вторых, повышается вероятность потери данных, так как число допустимых для отказа накопителей в массиве уже как минимум на 1 меньше. Например, RAID 5 в такой ситуации уже не будет защищен контрольными суммами, и повторный отказ будет иметь фатальные последствия. В-третьих, появляется шанс наткнуться на ошибку чтения с поверхности диска (bad block), так как пересчитываются все диски и происходит обращение к редко запрашиваемым данным.

В большинстве случаев массив продолжает свою работу во время ребилда, поэтому для восстановления данных выставляется приоритет нагрузки, в среднем от 10 до 30 процентов вычислительной мощности<sup>1</sup>. Чем ниже будет выставлен приоритет и чем больше емкость накопителей, тем длительнее будет процесс восстановления. В разных ситуациях он может занимать от нескольких часов до нескольких дней.

Благодаря быстрому расчету контрольных сумм RAID-массив в RAIDIX позволяет значительно сократить время ребилда по сравнению с существующими на рынке решениями.

<sup>1</sup> Выставление приоритета реконструкции также зависит от технических возможностей, заложенных производителем.

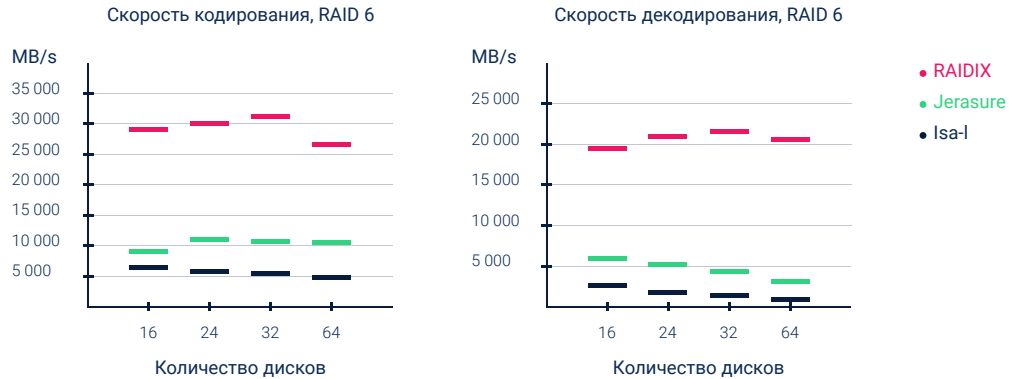
Сравнение длительности ребилда RAID 6 программного массива RAIDIX и существующего аппаратного RAID-контроллера для накопителей из 15 и 30 накопителей. Приоритет реконструкции 15%, объем накопителей 10 ТБ.



## Уникальный механизм вычислений

Механизм быстрой работы программного массива RAIDIX основан на оригинальном подходе к векторизации вычислений с расширениями SSE и AVX процессоров Intel при использовании RS-кодов. Ключевая идея этого подхода заключается в особом размещении данных в векторных регистрах, при котором существенно увеличивается скорость кодирования и декодирования данных.

Сравнение скорости кодирования и декодирования на примере RAID 6. Значения RAIDIX сравниваются с результатами библиотек ISA-L (Intel) и Jerasure.



## Возможности программного RAID-массива

Способность быстро кодировать и декодировать данные позволяет нашему RAID-массиву обеспечивать стабильный уровень производительности, необходимый для бесперебойной работы бизнес-приложений. С этой технологией отказ одного или нескольких накопителей не приводит к просадке производительности и не прерывает выполнение текущих задач.

Особое значение это имеет для высоконагруженных систем и инфраструктур с большими объемами хранения, где отказ даже одного накопителя подразумевает пересчет контрольных сумм для большого количества данных.

В RAIDIX был разработан целый ряд технологий, использующих преимущества быстрого RAID-массива и позволяющих заметно повысить пользовательские характеристики СХД.

### RAID 7.3

RAID-массив с тремя контрольными суммами практически не уступает по производительности RAID 6 на последовательных нагрузках, но при этом демонстрирует колоссальную надежность.

### RAID N+M

Возможность создавать RAID-массив с самостоятельным назначением количества накопителей (до 32) под контрольные суммы.

### Упреждающая реконструкция

Возможность быстро читать данные из контрольных сумм для сохранения скорости работы массива, исключая из процесса чтения наиболее медленные накопители.

### Частичная реконструкция

Сокращение времени на восстановление данных массива при случайном или запланированном извлечении накопителя.

### Защита от скрытого повреждения данных

Благодаря высокой скорости массива система может сканировать целостность данных при помощи контрольных сумм практически без снижения производительности.