

РЕАЛИЗАЦИЯ ПРОГРАММНОГО RAID

В RAIDIX программный RAID реализован на уровне ядра операционной системы. Он разрабатывался специально для обеспечения высокой производительности и доступности данных.

Благодаря уникальному методу расчета контрольных сумм, использующему оригинальное размещение обрабатываемых данных на векторных регистрах процессора, удается достигать высочайшей скорости восстановления данных.



ПРИМЕНЕНИЕ

В системах хранения данных RAID-массивы используются как для создания виртуального пространства на основе имеющихся физических носителей, так и для сохранности информации при выходе дисков из строя. При этом главным фактором производительности RAID-массива является скорость вычисления контрольных сумм и скорость восстановления данных из этих контрольных сумм.

Актуальным требованием для СХД является баланс между скоростью восстановления данных RAID-массива, количеством избыточных дисков и уровнем доступности данных при отказах.

В RAIDIX реализована уникальная модель программного массива, которая обладает рекордной скоростью восстановления данных в сочетании с высоким уровнем надежности и доступности.

ОСОБЕННОСТИ РАСЧЕТНОГО МОДУЛЯ RAID

Расчет контрольных сумм в программных RAID-массивах основан на векторизации вычислений с использованием расширений SSE4.2 и AVX процессоров Intel.

При этом, ключевое преимущество RAIDIX состоит в оригинальном подходе к векторным вычислениям, который позволил увеличить скорость кодирования и декодирования данных более чем в два раза по сравнению с другими, уже «разогнанными» векторизацией, библиотеками (ISA-L и Jerasure).

25 GBps

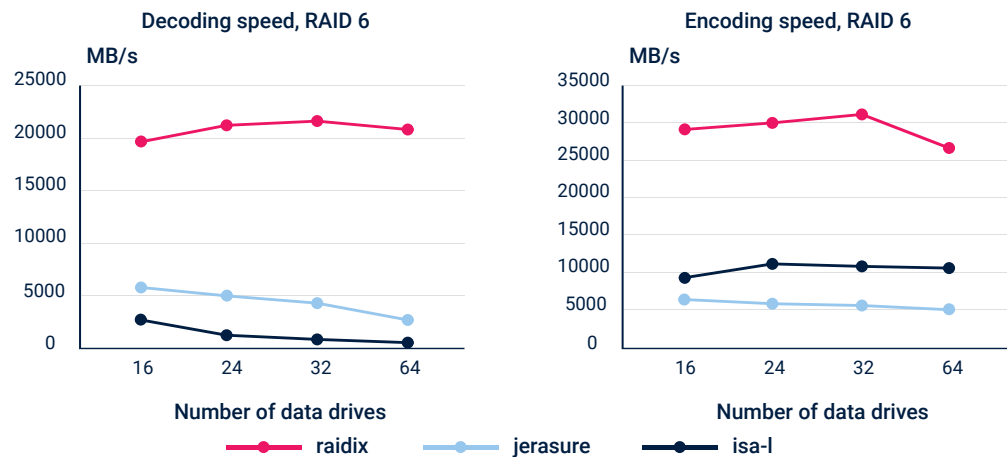
Скорость расчета контрольных сумм (на 1 вычислительное ядро)

Не более 10%

Просадка производительности при отказах на последовательных запросах

Сравнение скорости кодирования и декодирования на примере RAID 6.

Значения RAIDIX сравниваются с результатами библиотек ISA-L (Intel) и Jerasure.



Во многих системах хранения данных, использующих технологию RAID, расчёт контрольных сумм производится кодами Рида-Соломона с помощью арифметики конечных полей Галуа. Благодаря особому расположению данных в векторных регистрах, RAIDIX позволяют выполнять операцию умножения в поле сразу над 512 элементами, исполнив несколько простых инструкций.

Стоит отметить, что этот метод расчетов прекрасно параллелизуется, и суммарная производительность алгоритма растет почти линейно в многоядерных и многосоккетных системах.

СОБСТВЕННЫЕ УРОВНИ RAID

Важной особенностью реализации RAID-массивов в RAIDIX является наличие собственных уникальных моделей: RAID 7.3 и RAID N+M.

RAID 7.3

RAID 7.3 – уровень чередования блоков с тройным распределением четности, который позволяет восстанавливать данные при отказе до 3-х дисков. RAID 7.3 является аналогом RAID 6, но имеет более высокую степень надёжности, благодаря расчету сразу трех контрольных сумм. Для хранения контрольных сумм отводится емкость трех дисков.

RAID 7.3 существенно снижает вероятность отказа дисков без потерь в производительности и стоимости, поэтому рекомендуется использовать этот уровень для массивов объемом более 32 ТБ.

В RAID 7.3 для хранения контрольных сумм отводится емкость 3-х дисков.

D _{1,1}	D _{1,2}	D _{1,3}	S _{1,4}	S _{1,5}	S _{1,6}
D _{2,1}	D _{2,2}	S _{2,3}	S _{2,4}	S _{2,5}	D _{2,6}
D _{3,1}	S _{3,2}	S _{3,3}	S _{3,4}	D _{3,5}	D _{3,6}
S _{4,1}	S _{4,2}	S _{4,3}	D _{4,4}	D _{4,5}	D _{4,6}
S _{5,1}	S _{5,2}	D _{5,3}	D _{5,4}	D _{5,5}	S _{5,6}
S _{6,1}	S _{6,2}	D _{6,3}	D _{6,4}	S _{6,5}	S _{6,6}

RAID N+M

RAID N+M – уровень чередования блоков с M распределением четности, основанный на проприетарном алгоритме RAIDIX. RAID N+M позволяет пользователю самостоятельно определить количество дисков, выделяемых под хранение контрольных сумм. Это обеспечивает высочайший уровень надежности – уникальная технология RAIDIX позволяет восстановить данные при отказе до 32 дисков (в зависимости от количества дисков, выделяемых под контрольные суммы).

В RAID N+M M дисков выделяется под контрольные суммы.

N обозначает количество дисков под хранимые данные.

N						M			
D _{1,1}	D _{1,2}	D _{1,3}	D _{1,4}	...	D _{1,N}	S _{1,1}	S _{1,2}	...	S _{1,M}
D _{2,1}	D _{2,2}	D _{2,3}	...	D _{2,N-1}	S _{2,N}	S _{2,1}	...	S _{2,M-1}	D _{2,N}
D _{3,1}	D _{3,2}	...	D _{3,N-2}	S _{3,N-1}	S _{3,N}	...	S _{3,M-2}	D _{3,N-1}	D _{3,N}
D _{4,1}	...	D _{4,N-3}	S _{4,N-2}	S _{4,N-1}	...	S _{4,M-3}	D _{4,N-2}	D _{4,N-1}	D _{4,N}
...	D _{5,N-4}	S _{5,N-3}	S _{5,N-2}	...	S _{5,M-4}	D _{5,N-3}	D _{5,N-2}	D _{5,N-1}	D _{5,N}
D _{6,N-5}	S _{6,N-4}	S _{6,N-3}	...	S _{6,M-5}	D _{6,N-4}	D _{6,N-3}	D _{6,N-2}	D _{6,N-1}	D _{6,N}
S _{7,N-5}	S _{7,N-4}	...	S _{7,M-6}	D _{7,N-5}	D _{7,N-4}	D _{7,N-3}	D _{7,N-2}	D _{7,N-1}	D _{7,N}



ВОЗМОЖНОСТИ ТЕХНОЛОГИИ

Реализация программного массива в RAIDIX является ключевой технологией продукта. Глубокие математические исследования позволили создать действительно мощный инструмент, который сочетает в себе производительность, экономичность и высокий уровень сохранности данных.

Данная технология характеризуется рядом ключевых особенностей:

- Рекордная скорость расчета контрольных сумм, многократно превосходящая существующие аналоги.
- Возможность самостоятельного определения количества дисков для контрольных сумм (для RAID N+M).
- Повышение эффективности сопутствующих технологий – упреждающей реконструкции и частичной реконструкции.
- Высокий уровень отказоустойчивости при использовании массивов RAID 7.3 и RAID N+M.