

XII международная конференция
CEE-SECR / РАЗРАБОТКА ПО

28 - 29 октября, Москва



Облачный Росомаха

Евгений Анастасиев



О чем речь?

- Блочные СХД
- Локальные и распределенные



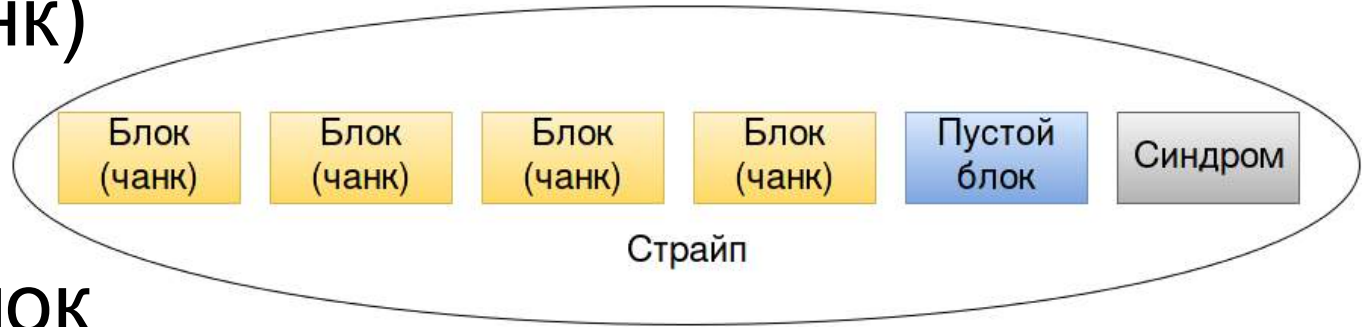
Какие проблемы есть

- Масштабируемость
- Быстродействие
- Отказоустойчивость



Важные термины

- Блок (чанк)
- Страйп
- Синдром
- Empty-блок

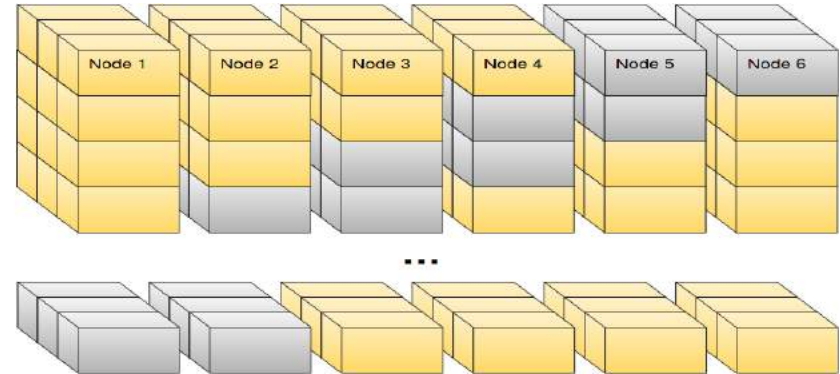


Отказоустойчивость в системах из нескольких компонент

RAID

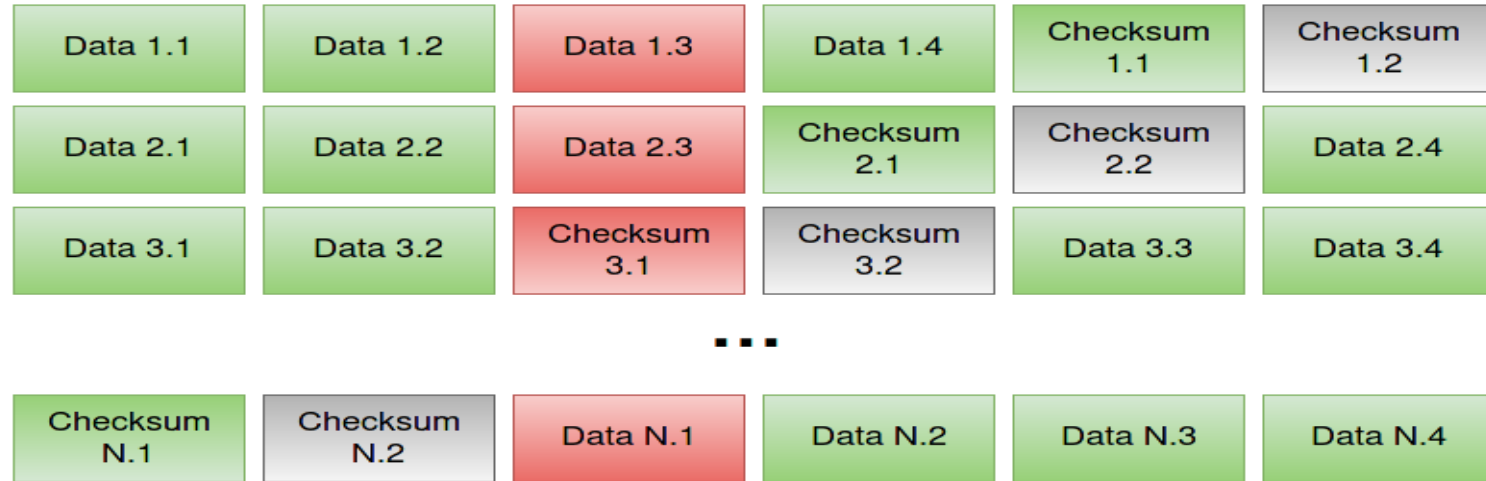


Кластер



Восстановление из контрольных сумм

- Для восстановления нужно прочитать контрольные суммы



- Меньше чтений — больше скорость восстановления



LRC и Regenerating Codes

Local Reconstruction Codes

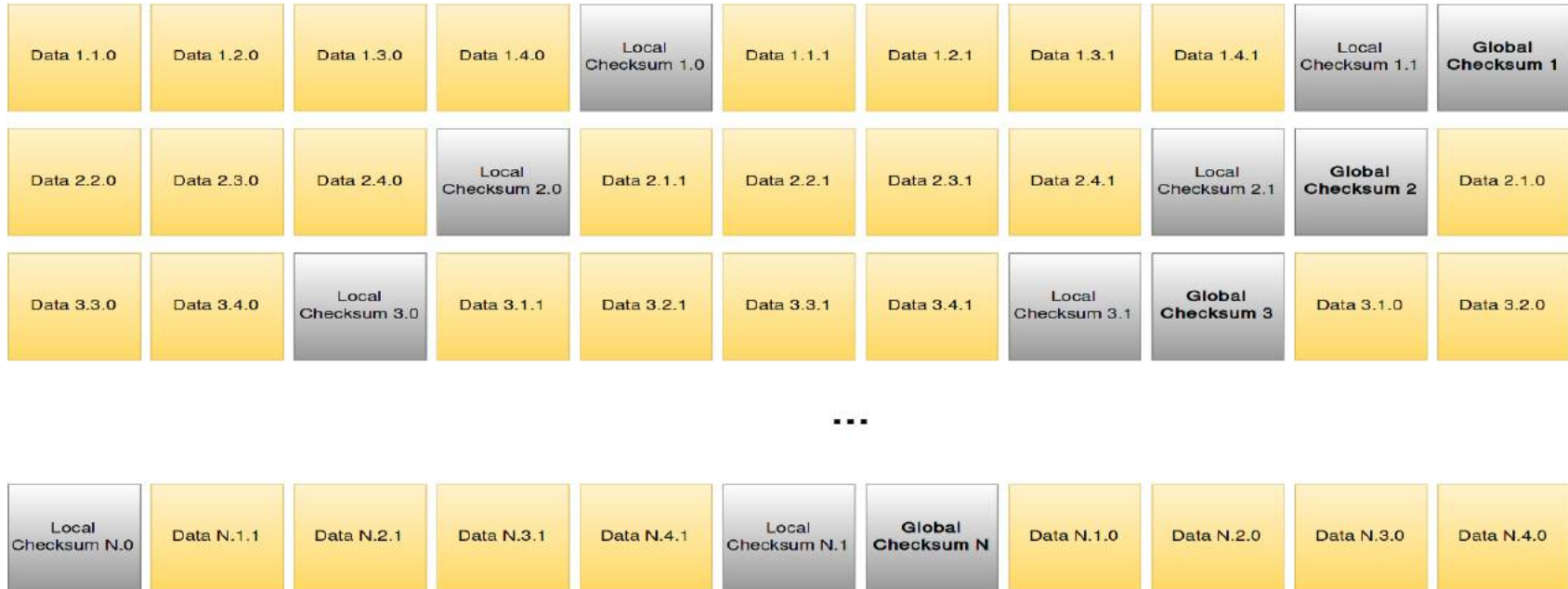
- Деление на локальные группы с контролем четности
- Мало чтений
- Повышенная избыточность

Regenerating Codes

- XOR-кодирование
- Минимальное количество чтений
- Нестандартный расчет синдромов



LRC

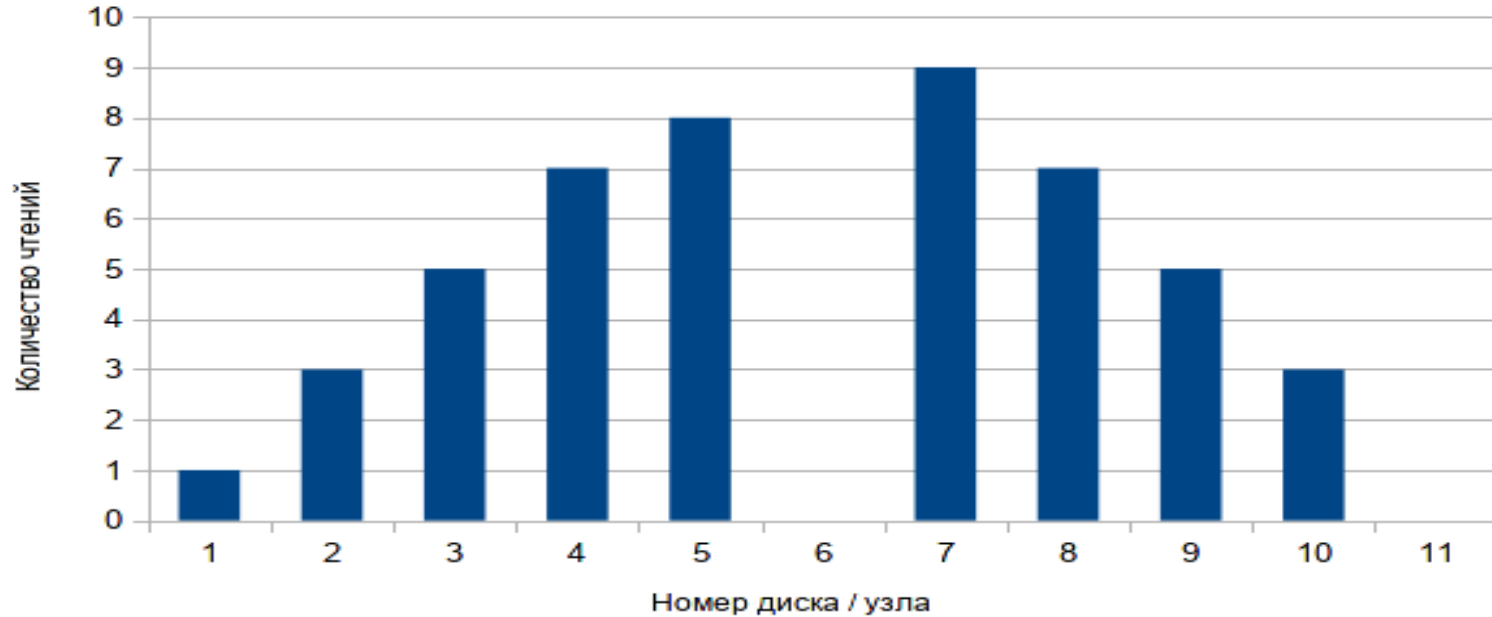


LRC: проблема «колокола»

0	0	0	0	S ₀	1	1	1	1	S ₁	G
0	0	0	S ₀	1	1	1	1	S ₁	G	0
0	0	S ₀	1	1	1	1	S ₁	G	0	0
0	S ₀	1	1	1	1	S ₁	G	0	0	0
S ₀	1	1	1	1	S ₁	G	0	0	0	0
1	1	1	1	S ₁	G	0	0	0	0	S ₀
1	1	1	S ₁	G	0	0	0	0	S ₀	1
1	1	S ₁	G	0	0	0	0	S ₀	1	1
1	S ₁	G	0	0	0	0	S ₀	1	1	1
S ₁	G	0	0	0	0	S ₀	1	1	1	1
G	0	0	0	0	S ₀	1	1	1	1	S ₁
1	3	5	7	8	0	9	7	5	3	0



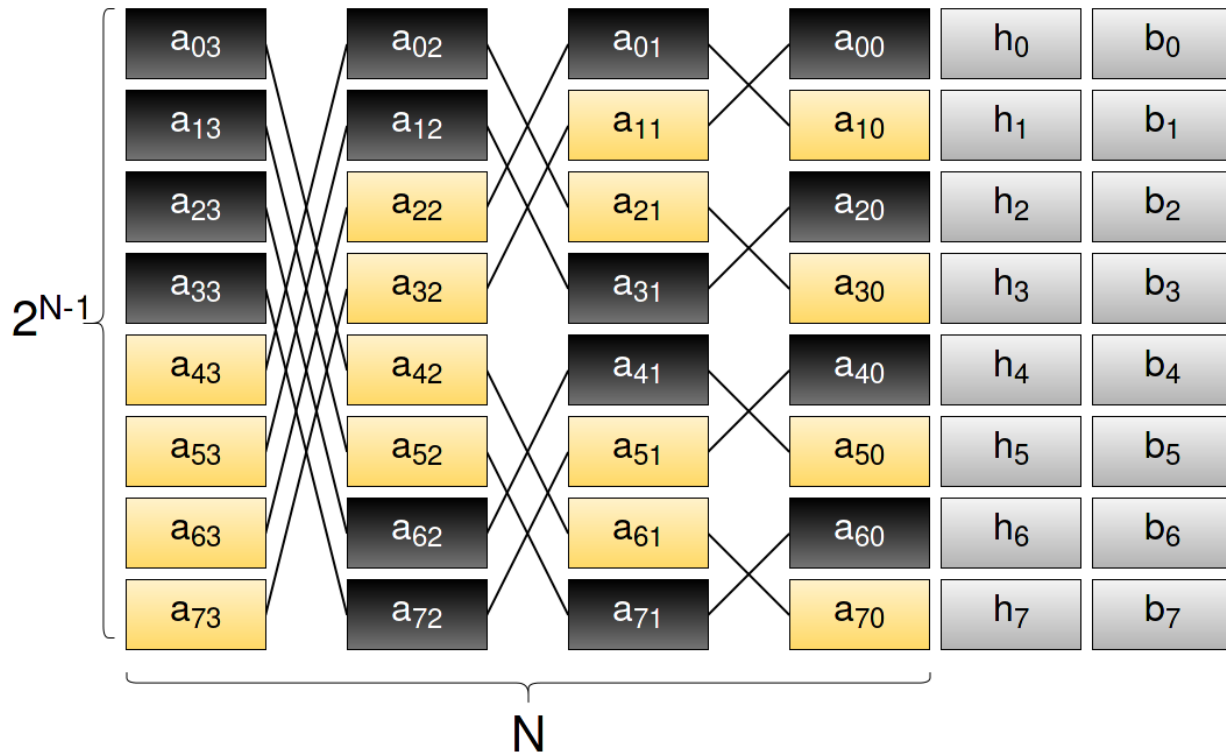
LRC: проблема «колокола»



Butterfly-схема

$$h_i = \sum_k a_{ik}$$

$i_2[j] = i_2[j-1] \Rightarrow a_{ij} \text{ — black}$

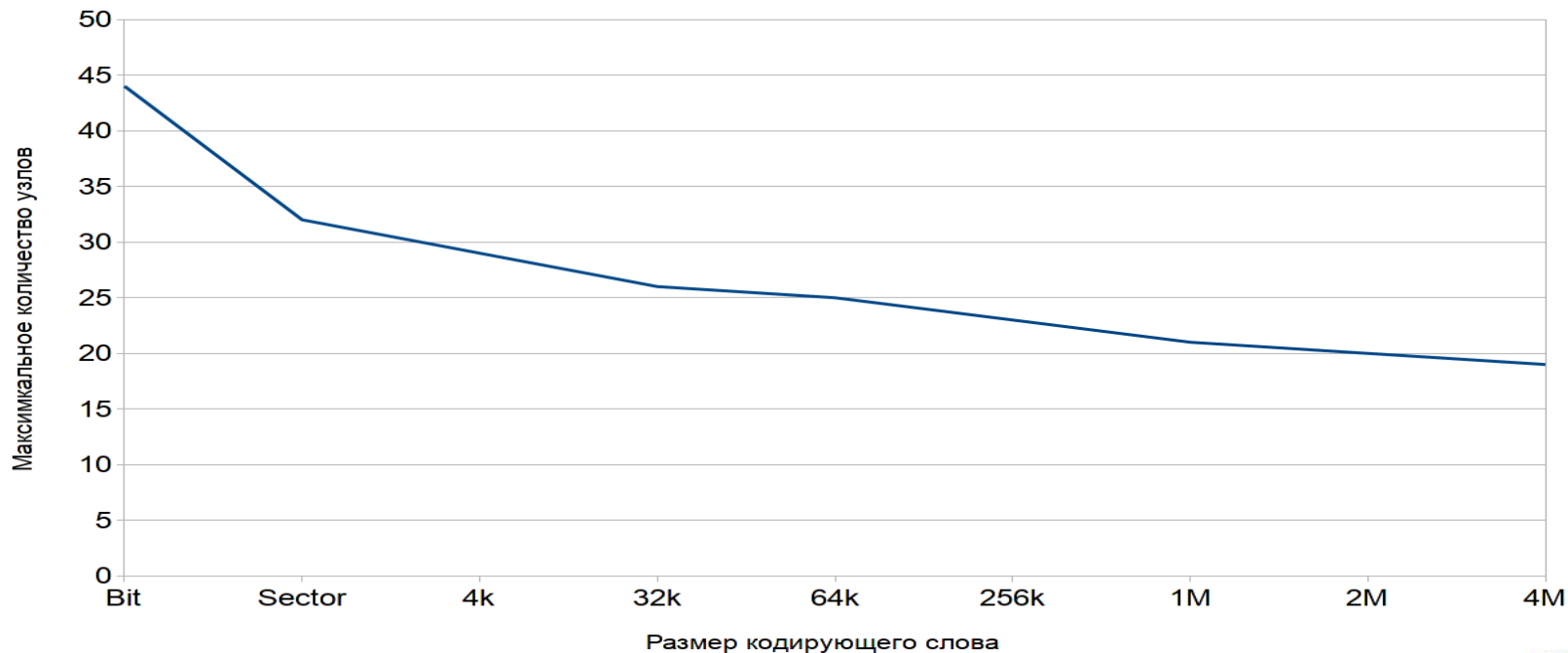


Butterfly-схема: проблема масштабирования

- Строгое ограничение по соотношению узлов и кодирующих символов
- Нужно выбирать между:
 - Расширением малым числом дисков
 - Скоростью работы
 - Максимальным размером кластера



Butterfly-схема: проблема масштабирования



Рандомизация в LRC

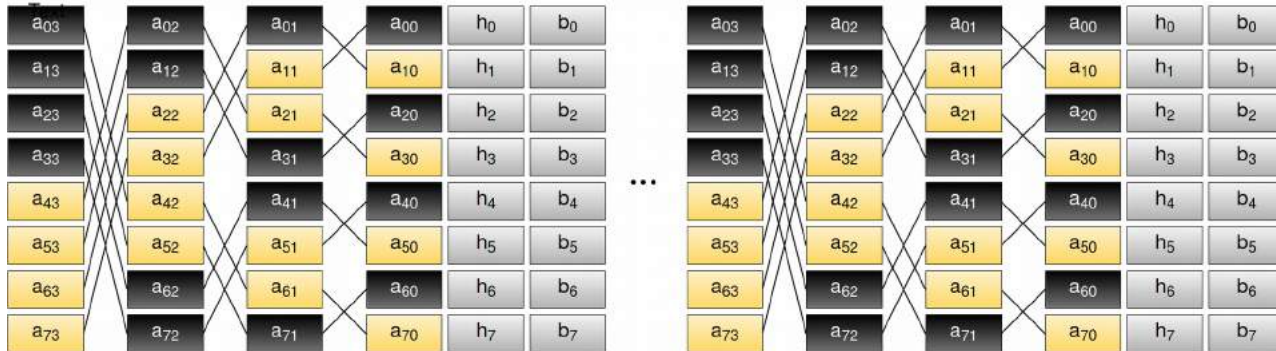
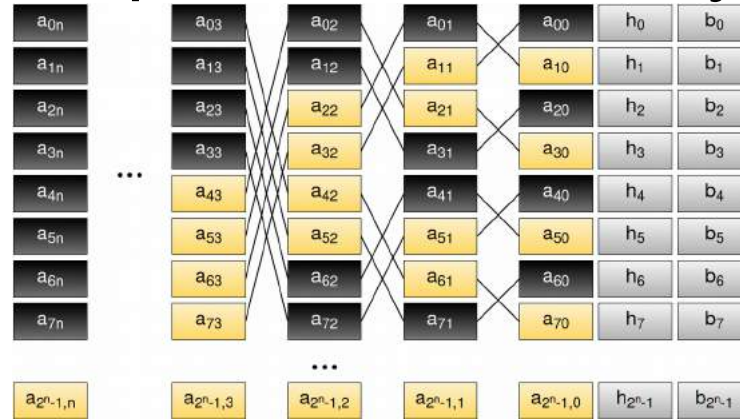
1	0	S ₀	0	1	1	G	E	1	S ₁	0	0
G	S ₁	0	0	1	E	0	1	0	S ₀	1	1
S ₀	0	1	1	0	0	1	G	1	0	E	S ₁
1	0	1	S ₁	G	0	S ₀	0	E	0	1	1
G	1	1	0	0	0	S ₀	1	1	E	S ₁	0
0	S ₀	0	0	1	S ₁	0	1	1	1	G	E
1	1	0	G	1	S ₁	1	0	0	0	E	S ₀
1	1	0	0	S ₀	0	E	1	S ₁	0	1	G
1	0	G	1	1	0	1	S ₀	0	S ₁	E	0
E	0	1	S ₁	0	1	0	1	0	1	S ₀	G
0	1	0	S ₀	E	1	G	S ₁	0	1	0	1
0	G	1	1	0	1	0	0	1	S ₀	S ₁	E
2	7	4	2	4	6	3	0	6	7	3	4

Рандомизация: решение проблемы «колокола»

- Чтение блока с каждого диска равновероятно
- Уже на 100000 страйпов разность максимума и минимума менее 1%
- Проблемы «колокола» не возникает



Применение локальных групп для масштабирования Butterfly-схемы

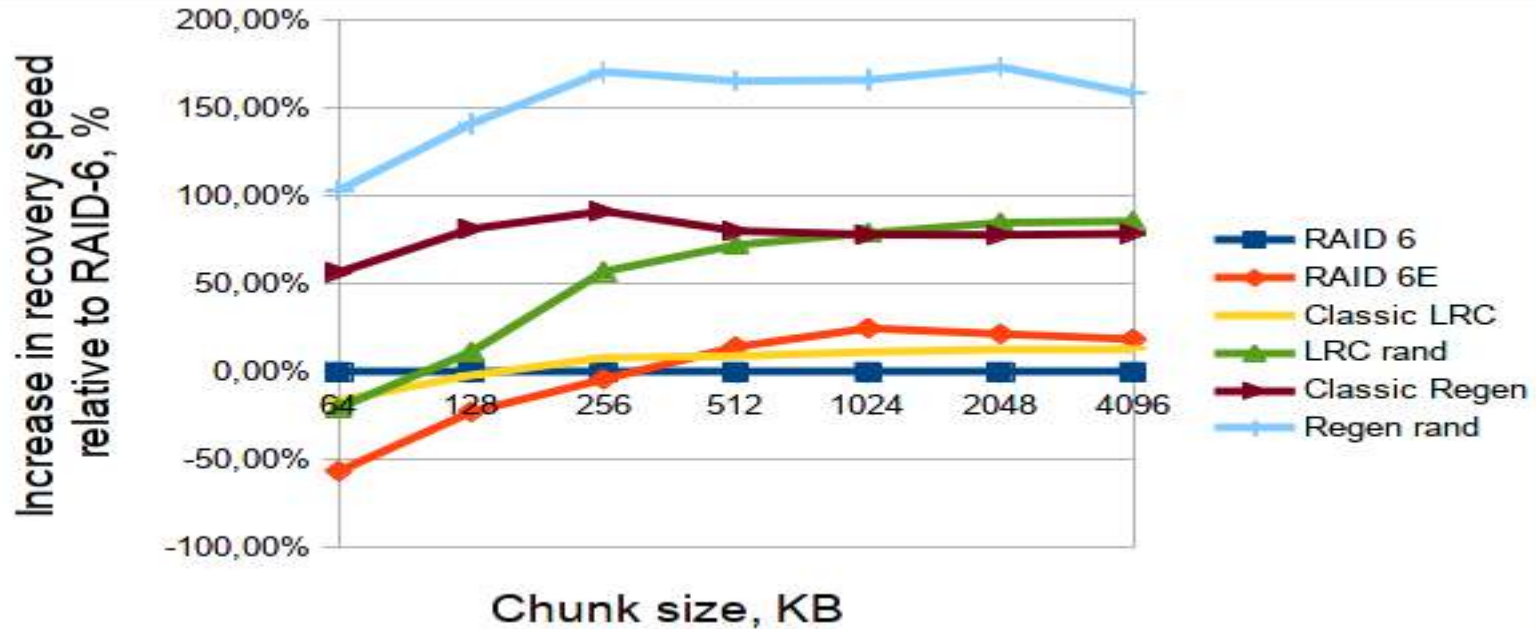


Рандомизация в Butterfly-схеме

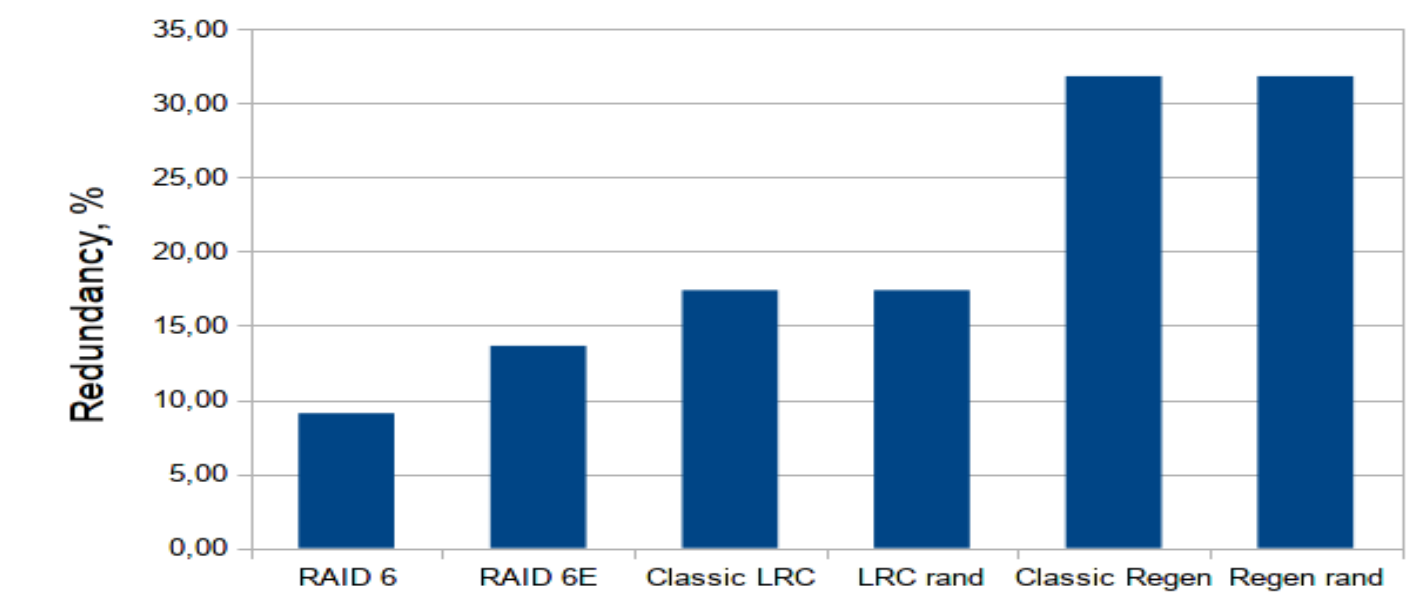
- Дает те же преимущества, что и в LRC
- Не только минимальное количество чтений, но и равномерное их распределение



Результаты: производительность



Результаты: избыточность





Спасибо за внимание!

www.raidix.ru
request@raidix.com