

**ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ**  
**Федеральное государственное образовательное бюджетное**  
**учреждение высшего профессионального образования**  
**«САНКТ-ПЕТЕРБУРГСКИЙ**  
**ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ**  
**им. проф. М. А. БОНЧ-БРУЕВИЧА»**

---

**А.Н.Губин**

## **Сети хранения данных**

**Краткий конспект лекций**

**Раздел 4. Защита данных. RAID-массивы**

**САНКТ-ПЕТЕРБУРГ**  
**2017**

## **Содержание**

### **Раздел 1. Содержание курса, цели и задачи дисциплины**

- 1.1. Общее содержание курса, цели и задачи изучения дисциплины.*
- 1.2. Структура дисциплины и ее связь с другими курсами.*
- 1.3. Хранение информации как основная задача информационных центров.*

### **Раздел 2. Общие характеристики процессов хранения и управления данными**

- 2.1. Основные технологии хранения данных.*
- 2.2. Структурированные и неструктурированные данные.*
- 2.3. Общие характеристики процессов хранения и управления данными.*
- 2.4. Общая структура информационных центров.*
- 2.5. Управление хранением данных. Многоуровневое хранение данных.*

### **Раздел 3. Основные компоненты среды хранения данных**

- 3.1. Физические компоненты. Хост, дисковое устройство.*
- 3.2. Логические компоненты. Файловые системы.*
- 3.3. Производительность дисковых устройств.*

### **Раздел 4. Защита данных. RAID-массивы**

- 4.1. Распределение данных. Зеркалирование данных.*
- 4.2. Контроль четности.*
- 4.3. Основные конфигурации RAID-массивов.*
- 4.4. IOPS-операций и конфигурация дисков.*
- 4.5. Влияние наличия RAID на производительность дисковых устройств.*

### **Раздел 5. Системы хранения данных. Интеллектуализация систем хранения данных**

- 5.1. Основные компоненты интеллектуальных систем хранения информации.*
- 5.2. Операция чтения и записи данных с использованием КЭШ-памяти.*
- Защита КЭШ- данных*
- 5.3. Основные компоненты системы хранения данных Symmetrix.*

### **Раздел 6. Сети хранения данных**

- 6.1. Архитектура сетей хранения данных. Особенности архитектуры СХД Fiber- Channel. Зонирование данных.*
- 6.2. Контекстная адресация данных.*
- 6.3. Виртуализация систем хранения данных.*

### **Раздел 7. Перспективные направления развития технологий хранения данных и управления информацией**

- 7.1. Внеполостная и внутриполостная виртуализация систем данных.*
- 7.2. Основные проблемы виртуализации систем хранения данных.*

## Раздел 4. Защита данных. RAID-массивы

### 4.1. Распределение данных. Зеркалирование данных.

Неотъемлемой частью современных вычислительных систем являются жесткие магнитные диски (HDD). Надежность работы HDD характеризуется средним временем безотказной работы (это наработка на отказ). Современные информационные центры используют в своих системах хранения данных тысячи жестких дисков. Чем больше дисков, тем больше вероятность выхода из строя одного из используемых дисков. Так если СХД использует 100 дисков, то при среднем времени безотказной работы для каждого диска 750 000 часов, средний период безотказной работы всего массива дисков составит 7 500 часов (750 000/100). То есть в среднем через 7 500 часов (через каждые 10-11 месяцев) один из жестких дисков системы выходит из строя.

Кроме того, быстродействие HDD существенно меньше быстродействия остальных компонентов СХД.

RAID-уровни формируются с использованием следующих технологий:

- распределения;
- зеркалирования;
- контроля четности.

Эти технологии определяют доступность данных, производительность массива и надежность хранения данных. Некоторые RAID-массивы используют одну из перечисленных технологий, другие могут использовать одновременно несколько технологий. Выбор уровня RAID-массива определяется требованиями производительности приложения доступности и надежности хранения данных.

#### **RAID 0 (распределение)**

Минимальное количество дисков 2.

Объем массива = число дисков \* размер наименьшего диска.

При записи данные разбиваются на блоки и пишутся на каждый диск массива. Избыточность отсутствует, выход одного из дисков приводит к полной потере информации. Используется только для ускорения чтения/записи, применяется там, где не важна надежность, а нужна скорость и низкая стоимость. Например, как массив для виртуальных машин. При увеличении количества дисков в массиве увеличивается производительность, так как появляется возможность одновременного считывания и записи

большого количества блоков данных. Не рекомендуется для хранения важных данных, так как надежность массива в целом ниже, чем у отдельно взятого жесткого диска. При восстановлении такого массива главное выстроить образы дисков в правильном порядке. На рис. 4.1 представлена общая структура распределения данных по дисковому массиву при использовании RAID нулевого уровня.

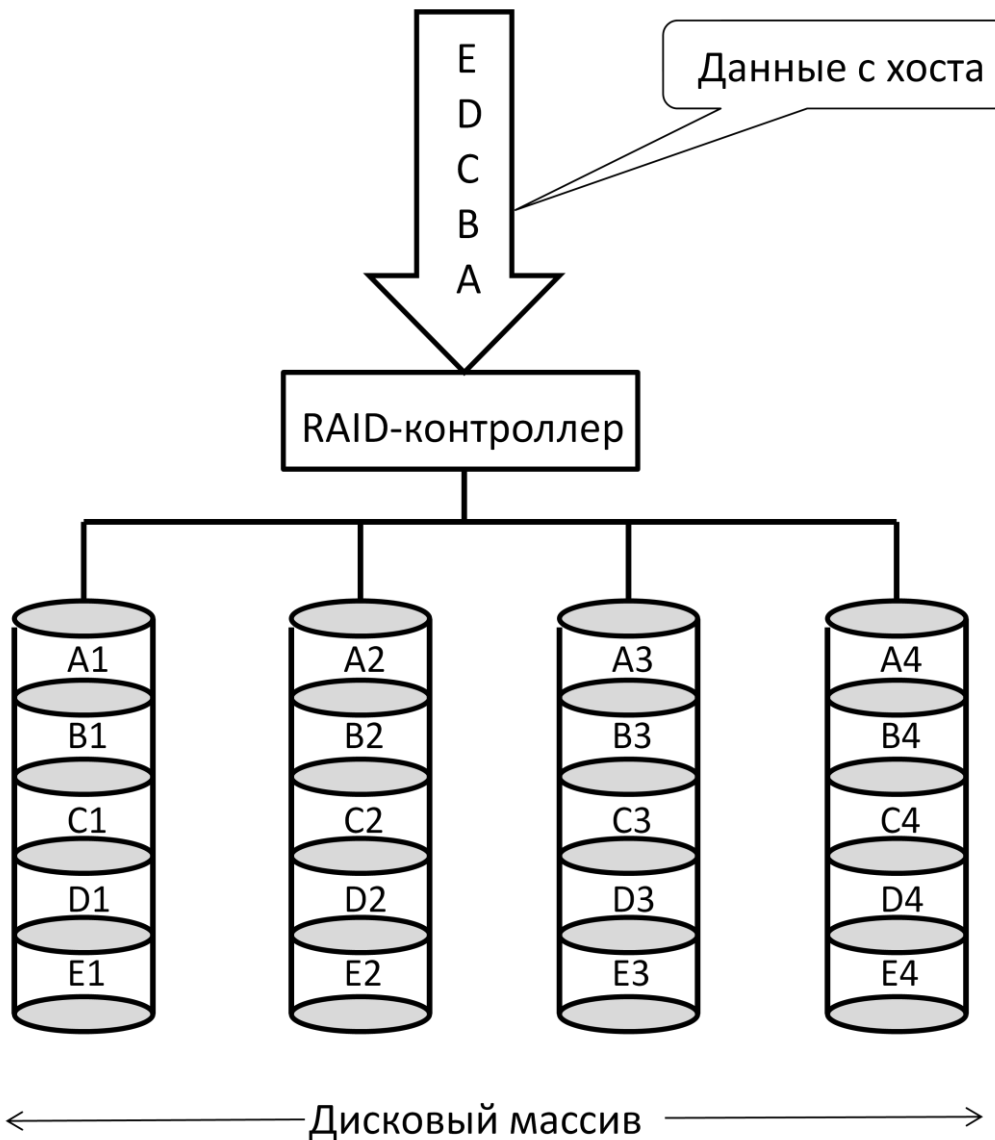


Рис. 4.1. Общая структура распределения данных при использовании RAID нулевого уровня

### **RAID 1 (зеркалирование)**

Минимальное количество дисков 2.

Объем массива = размер наименьшего диска.

Зеркалирование – это технология хранения данных на двух разных дисках (на втором диске хранится копия данных, записанных на первом диске). Производительность данного уровня RAID при записи равна производительности самого медленного диска в массиве, а вот при чтении

возможен выигрыш, читать можно одновременно со всех дисков. При этом это самый дорогой вариант RAID массива, так как дисковое пространство используется крайне не эффективно.

В случае сбоев в работе массива, восстановление данных осуществляется с минимальными по сравнению с другими RAID массивами затратами. RAID-контроллер использует для восстановления данных диск с зеркалированными данными. Конфигурация RAID 1 рекомендуется для использования с приложениями, требующими высокой надежности хранения данных. На рис. 4.2 представлена общая структура распределения данных по дисковому массиву при использовании RAID первого уровня.

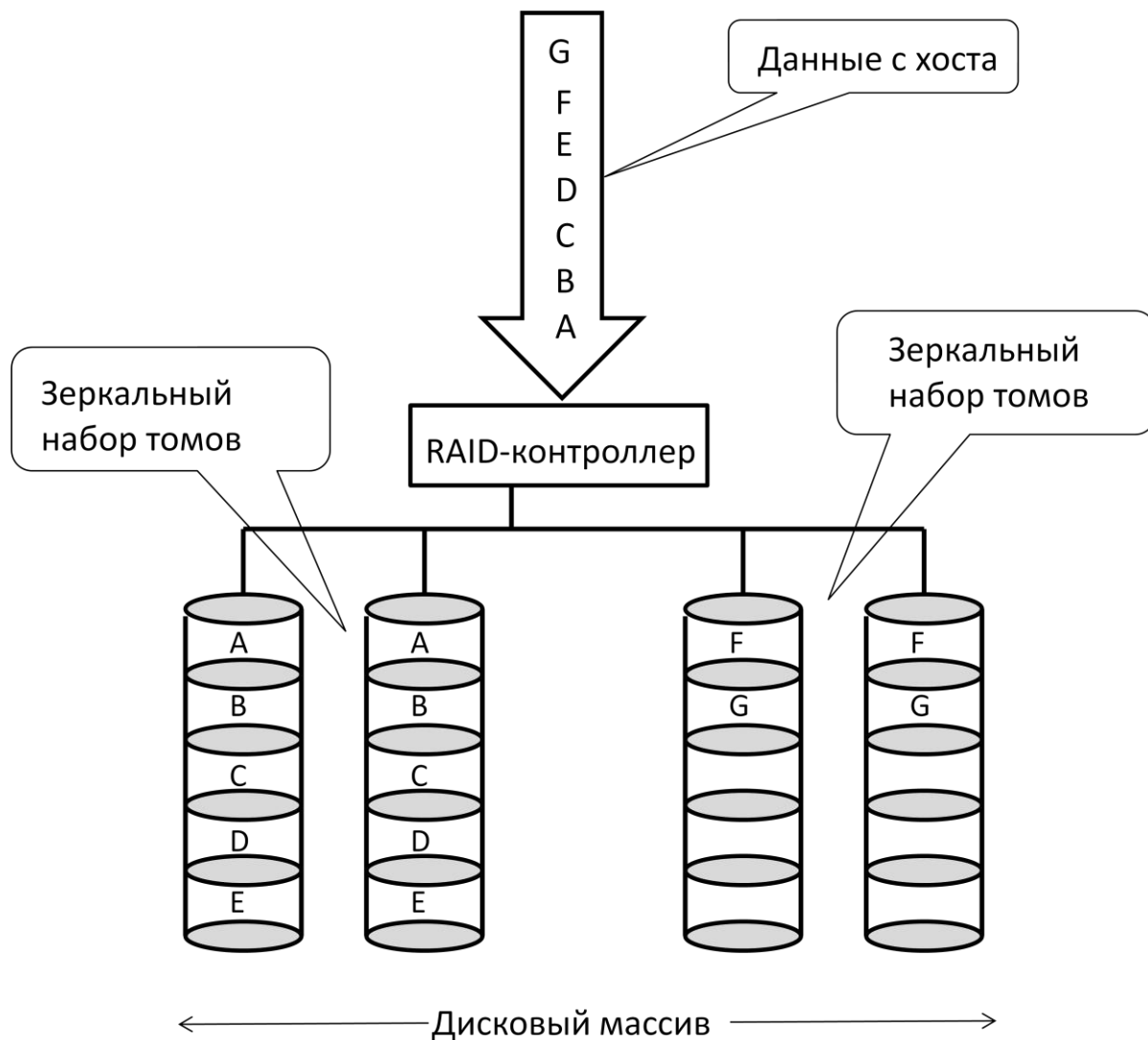


Рис. 4.2. Общая структура распределения данных при использовании RAID первого уровня

## **4.2. Контроль четности**

### **RAID 3 (контроль четности)**

RAID 3 распределяет данные по дисковому массиву для увеличения производительности и применяет контроль четности для повышения отказоустойчивости системы.

Структура массива RAID-3 такова: в массиве из  $n$  дисков данные разбиваются на блоки размером 1 байт и распределяются по  $n-1$  дискам, а еще один диск используется для хранения блоков четности. В RAID-2 для этой цели стояло  $n-1$  дисков, но большая часть информации на этих дисках использовалась только для коррекции ошибок на лету, а для простого восстановления в случае поломки диска, хватает и одного выделенного винчестера.

Соответственно, отличия RAID-3 от RAID-2 очевидны: невозможность коррекции ошибок на лету и меньшая избыточность. Преимущества таковы: скорость чтения и записи данных высока, а для создания массива требуется совсем немного дисков, всего три. Но массив этого типа хорош только для однозадачной работы с большими файлами, так как наблюдаются проблемы со скоростью при частых запросах данных небольшого объема. На рис. 4.3 представлена общая структура распределения данных по дисковому массиву при использовании RAID третьего уровня.

### **RAID 4 (контроль четности)**

RAID-4 похож на RAID-3, но отличается от него тем, что данные разбиваются на блоки, а не на байты. Таким образом, удалось устранить проблему низкой скорости передачи данных небольшого объема. Запись же производится медленно из-за того, что четность для блока генерируется при записи и записывается на единственный диск. Используются массивы такого типа очень редко.

Структура распределения данных по дисковому массиву при использовании RAID четвертого уровня не отличается от структуры, представленной на рис. 4.3.

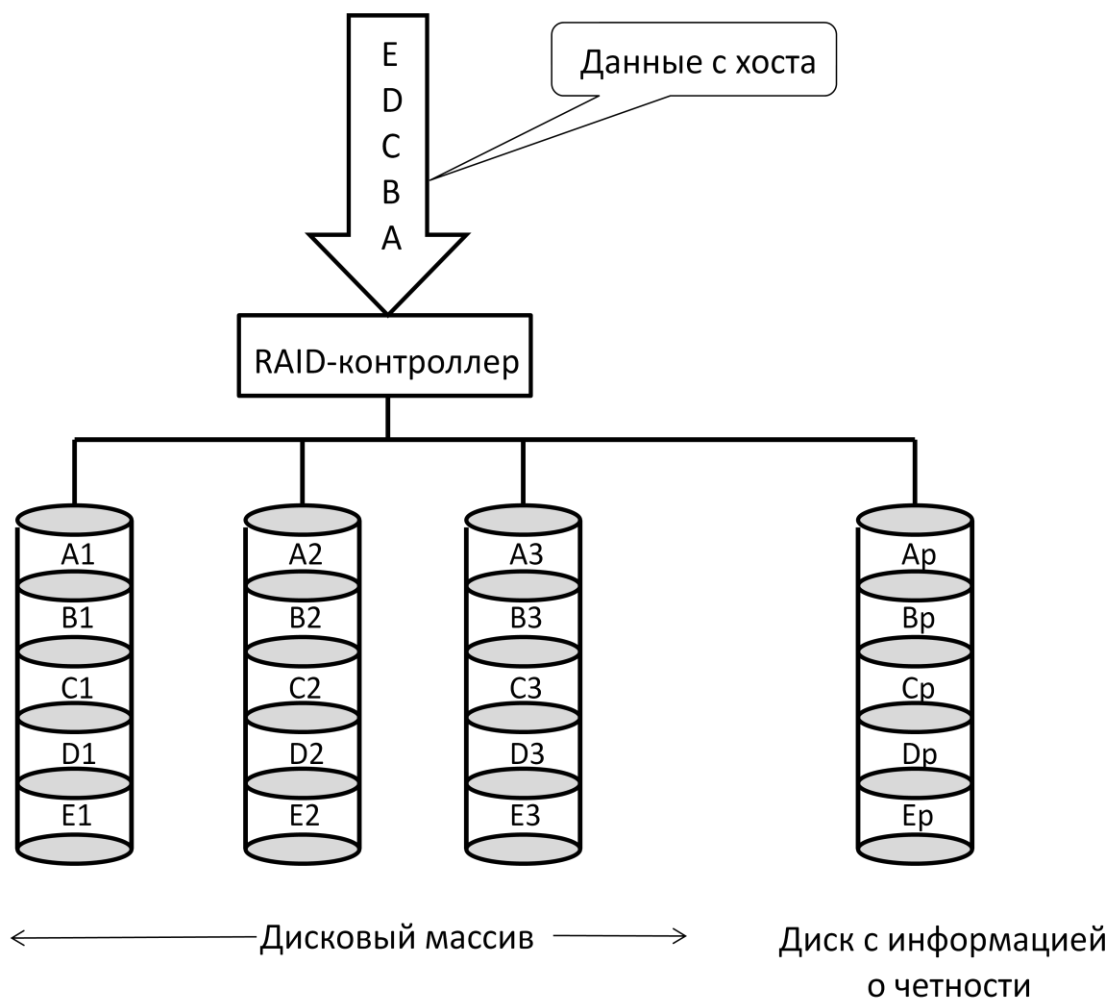


Рис. 4.3. Общая структура распределения данных при использовании RAID третьего уровня

### 4.3. Основные конфигурации RAID-массивов

Существует два типа реализации RAID: на аппаратном и программном уровнях.

Программная реализация RAID использует для управления информацией функции, реализованные на основе программ. Программное обеспечение RAID-технологий разрабатывается на уровне операционных систем без использования аппаратных контроллеров для управления RAID-массивов.

Использование программного RAID более рентабельно: не надо тратиться на приобретение дополнительных аппаратных средств, эксплуатация программных средств более просто и менее затратное по сравнению аппаратными средствами. Однако программный RAID обладает следующими недостатками:

- программная реализация RAID существенно снижает общую производительность информационной системы. Это происходит из-за

дополнительной нагрузки на процессор, который должен производить дополнительную работу по реализации функций RAID-массива;

- не все уровни RAID-массивов реализуются на программном уровне;
- программное обеспечение RAID связано с операционной системой хоста. Обновление программного обеспечения хоста вызывает необходимость проверки программного обеспечения на совместимость.

Аппаратная реализация RAID подразумевает разработку специализированного контроллера RAID-массива для хоста или для дискового массива. Тип реализации контроллера определяет способ взаимодействия дискового массива с хостом.

Плата RAID-контроллера устанавливается в хост, жесткие диски подключаются непосредственно к контроллеру, взаимодействие с жесткими дисками осуществляется через PCI-шины.

Внешний RAID-контроллер – это отдельное устройство, функционирующее как интерфейс между хостом и дисковым массивом. Основными функциями контроллера являются:

- управление группами дисков и контроль состояния дисков;
- перевод запросов на ввод/вывод между логическими и физическими дисками;
- восстановление данных при сбоях дискового массива.

Общая структура RAID-массива может быть представлена в следующем виде (рис. 4.4).

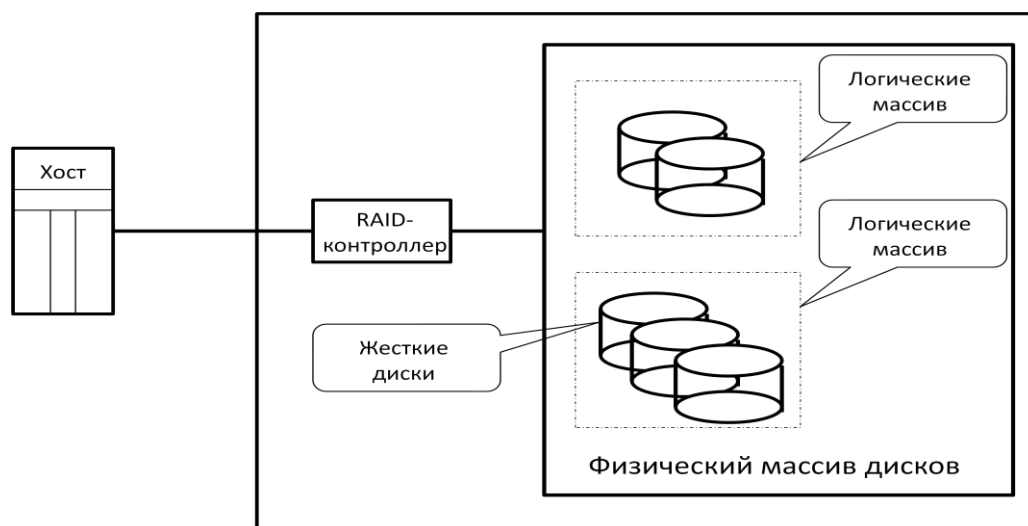


Рис. 4.4. Основные компоненты RAID-массива

Кроме указанных выше RAID 0, RAID 1, RAID 3 и RAID 4, для достижения целей обеспечения доступности данных, производительности дисковых массивов и надежности хранения данных используются следующие конфигурации RAID – массивов:

- RAID 2;
- RAID 5;



- RAID 6,  
а также различные их комбинации.

### **RAID 2 (коррекция ошибок)**

В массивах такого типа диски делятся на две группы - для данных и для кодов коррекции ошибок, причем если данные хранятся на  $n$  дисках, то для складирования кодов коррекции необходимо  $n-1$  дисков. Данные записываются на соответствующие винчестеры так же, как и в RAID-0, они разбиваются на небольшие блоки по числу дисков, предназначенных для хранения информации. Оставшиеся диски хранят коды коррекции ошибок, по которым в случае выхода какого-либо винчестера из строя возможно восстановление информации. В данной конфигурации надёжность достигается за счет применения кодов коррекции по алгоритму У. Хэмминг (Richard W. Hamming), который позволяет обнаруживать и исправлять битовые ошибки на лету (Error Detection and Correction). Однако держать ради этого громоздкую структуру из почти двойного количества дисков никому не хотелось, и этот вид RAID- массива не получил распространения.

### **RAID 5 (распределенное хранение информации о четности)**

Минимальное количество дисков 3.

Объем массива = (Число дисков - 1) \* Размер наименьшего диска.

Конфигурация RAID 5 похожа на конфигурацию RAID 4, поскольку использует сегментирование при независимом доступе к полосам. Разница заключается в размещении контрольных данных. В RAID 4 они записываются на отдельный диск. Этот диск становится узким местом при записи. В RAID 5 контрольная информация распределяется по всем дискам.

Под контрольной информацией подразумевается результат операции *xor* (исключающее или). Логическая операция *xor* обладает особенностью, которая даёт возможность заменить любой операнд результатом, и, применив алгоритм *xor*, получить в результате недостающий операнд. Например:  $a \text{ xor } b = c$  (где  $a$ ,  $b$ ,  $c$  — три диска рейд-массива), в случае если  $a$  откажет, можно получить информацию диска  $a$ , поставив на его место  $c$  и проведя *xor* между  $c$  и  $b$ :  $c \text{ xor } b = a$ . Это применимо вне зависимости от количества операндов:  $a \text{ xor } b \text{ xor } c \text{ xor } d = e$ . Если отказывает  $c$  тогда  $e$  встает на его место и проведя *xor* в результате получаем  $c$ :  $a \text{ xor } b \text{ xor } e \text{ xor } d = c$ . Этот метод, по сути, обеспечивает отказоустойчивость 5 версии. Для хранения результата *xor* требуется всего 1 диск, размер которого равен размеру любого другого диска в RAID.

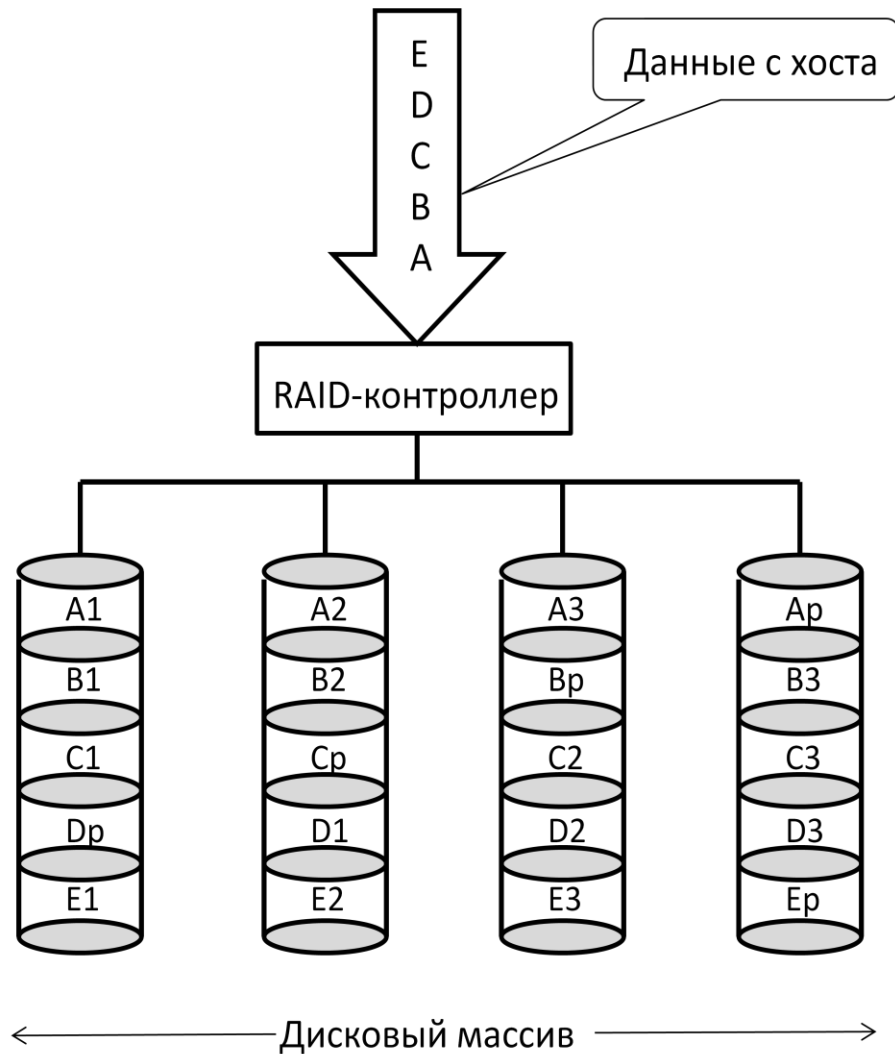


Рис. 4.5. Общая структура распределения данных при использовании RAID пятого уровня

На рис. 4.5 представлена общая структура распределения данных по дисковому массиву при использовании RAID пятого уровня. Места хранения контрольной информации обозначены как  $A_p$ ,  $B_p$ ,  $C_p$ ,  $D_p$ ,  $E_p$ .

Рассмотрим пример восстановления данных для RAID массива, структура которого изображена на рис.18. Имеется четыре диска для хранения данных и записи контрольных сумм. Необходимо записать на дисковый массив последовательность битов 1101 0011 1100. RAID- контролер разбивает эту последовательность на блоки по четыре бита, то есть  $A_1=1101$ ,  $A_2=0011$ ,  $A_3=1100$  и вычисляет контрольную сумму

$$A_p = A_1 \oplus A_2 \oplus A_3 = 1101 \oplus 0011 \oplus 1100 = 0010$$

Эта контрольная сумма, записываемая на четвертый диск.

Если один из дисков, например, третий, вышел из строя, то блок  $A_3 = 1101$  оказывается недоступным для считывания данных. Однако его значение легко восстановить по контрольной сумме и по значениям остальных блоков с помощью все той же операции «исключающего ИЛИ»:

$$A3=A1\oplus A2\oplus A_p=1101\oplus 0011\oplus 0010=1100$$

В случае RAID 5 все диски массива имеют одинаковый размер, однако общая емкость дисковой подсистемы, доступной для записи, становится меньше ровно на один диск. Например, если четыре диска имеют размер 100 Гбайт, то фактический размер массива составляет 300 Гбайт, поскольку 100 Гбайт отводится на контрольную информацию.

RAID 5 используется в основном для обслуживания систем средней производительности.

Недостатки RAID 5 проявляются при выходе из строя одного из дисков, при этом весь том переходит в критический режим (degrade), все операции записи и чтения сопровождаются дополнительными манипуляциями, резко падает производительность. Уровень надежности снижается до надежности RAID-0 с соответствующим количеством дисков (то есть в  $n$  раз ниже надежности одиночного диска). Если до полного восстановления массива произойдет выход из строя, или возникнет невосстановимая ошибка чтения хотя бы еще на одном диске, то массив разрушается, и данные на нем восстановлению обычными методами не подлежат.

Следует принять во внимание, что процесс RAID Reconstruction (восстановления данных RAID за счет избыточности), после выхода из строя диска вызывает интенсивную нагрузку чтения с дисков на протяжении многих часов непрерывно, что может спровоцировать выход какого-либо из оставшихся дисков из строя в этот наименее защищенный период работы RAID, а также выявить ранее необнаруженные сбои чтения в массивах cold data (данных, к которым не обращаются при обычной работе массива, архивные и малоактивные данные), что резко повышает риск сбоя при восстановлении данных.

### **RAID 6 (распределенное хранение двух контрольных сумм)**

Минимальное количество дисков 4.

Объем массива = (Число дисков - 2) \* Размер наименьшего диска.

RAID 6 - похож на RAID 5, но имеет более высокую степень надёжности. Под контрольные суммы выделяется ёмкость 2-х дисков, рассчитываются 2 суммы по разным алгоритмам (XOR и алгоритм Рида-Соломона). Требуется более мощный RAID-контроллер. Обеспечивается работоспособность после одновременного выхода из строя двух дисков, то есть, реализована защита от кратного отказа. Для организации массива требуется минимум 4 диска. Обычно использование RAID-6 вызывает примерно 10-15% падение производительности дисковой группы, относительно RAID 5, что вызвано большим объёмом обработки для контроллера (необходимость рассчитывать вторую контрольную сумму, а также читать и перезаписывать больше дисковых блоков при записи каждого блока).

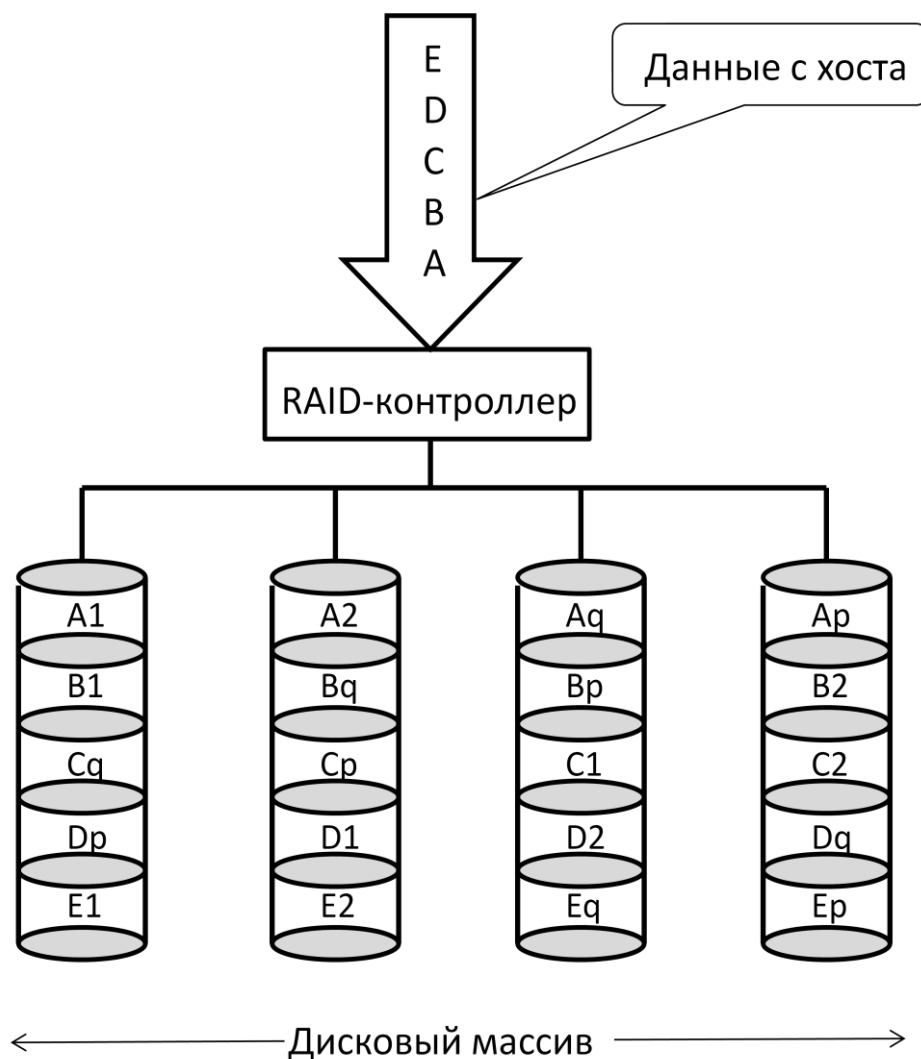


Рис. 4.6. Общая структура распределения данных при использовании RAID шестого уровня

Из-за своей излишней избыточности и потери производительности не получил широкого распространения и может быть рекомендован к использованию только при повышенных требованиях к надёжности.

Общая структура распределения данных по дисковому массиву при использовании RAID шестого уровня представлена на рис. 4.6. Места хранения контрольной информации обозначены как  $A_p, B_p, C_p, D_p, E_p$  – рассчитанные по *xor*-алгоритму и  $A_q, B_q, C_q, D_q, E_q$  – по алгоритму Рида-Соломона (недвоичные циклические коды, позволяющие исправлять ошибки в блоках данных).

### Комбинированные уровни RAID-массивов

Помимо стандартных («Common RAID Disk Drive Format (DEF) standard») базовых уровней RAID 0 - RAID 6, существуют комбинированные уровни с названиями вида «RAID  $\alpha+\beta$ » или «RAID  $\alpha\beta$ », что обычно означает «RAID  $\beta$  составленный из нескольких RAID  $\alpha$ ».

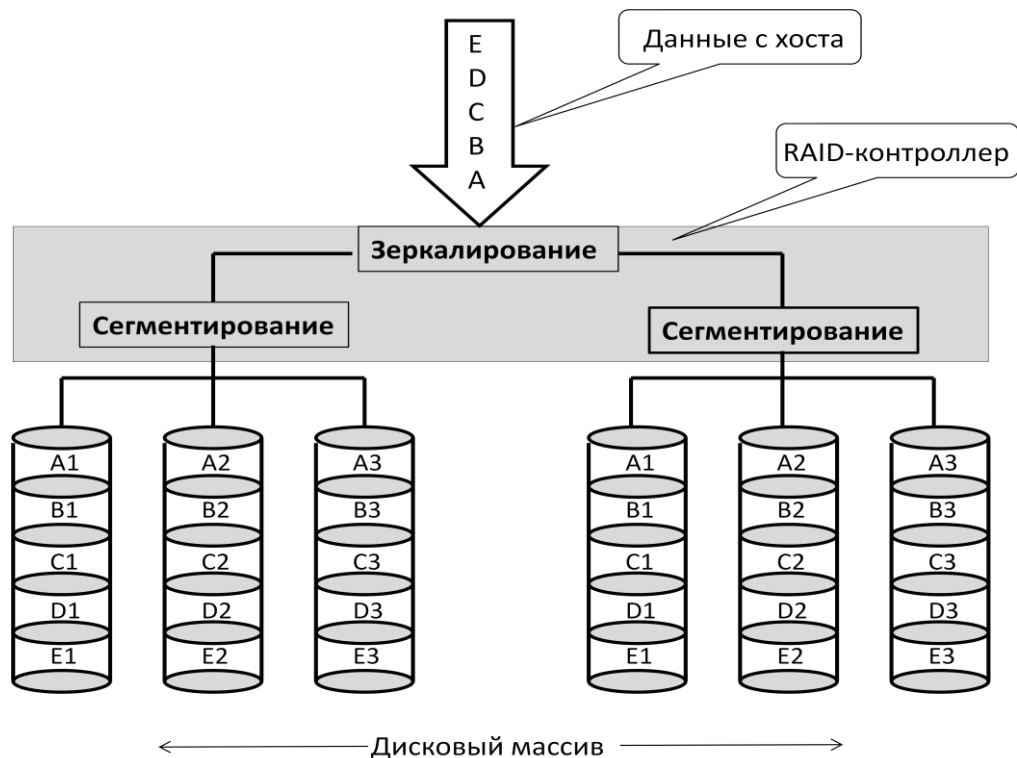


Рис. 4.7. Общая структура распределения данных при использовании RAID 10 (1+0)

Например:

- RAID 10 (или 1+0) — это RAID 0, составленный из нескольких (или хотя бы двух) RAID 1 (зеркалированных пар). Общая структура организации данных на дисках RAID 10 представлена на рис. 4.7.

- RAID 01 (или 0+1) — это RAID 1, составленный из нескольких (или хотя бы двух) RAID 0 (зеркалированных пар). Общая структура организации данных на дисках RAID 01 представлена на рис. 4.8.

- RAID 51 — RAID 1, зеркалирующий два RAID 5 .

При обычных условиях RAID 10 и RAID 01 обладают одинаковыми преимуществами. Различие между ними проявляется при операциях восстановления информации в случае выхода из строя одного из дисковых устройств массива. RAID 10 – сегментированное зеркало. Основным элементом RAID 10 является зеркальная пара. Это значит, что данные сначала зеркалируются, а потом обе копии данных распределяются по разным дискам в RAID 0. Если диск в любом задействованном зеркальном комплекте откажет, то его содержимое может быть получено с оставшегося диска зеркальной пары.

Таким образом, если массив RAID 10 подвергнется максимальному количеству отказов дисков, которое он может перенести, он преобразится в массив RAID 0, не имеющий надежности, но зато очень быстрый.

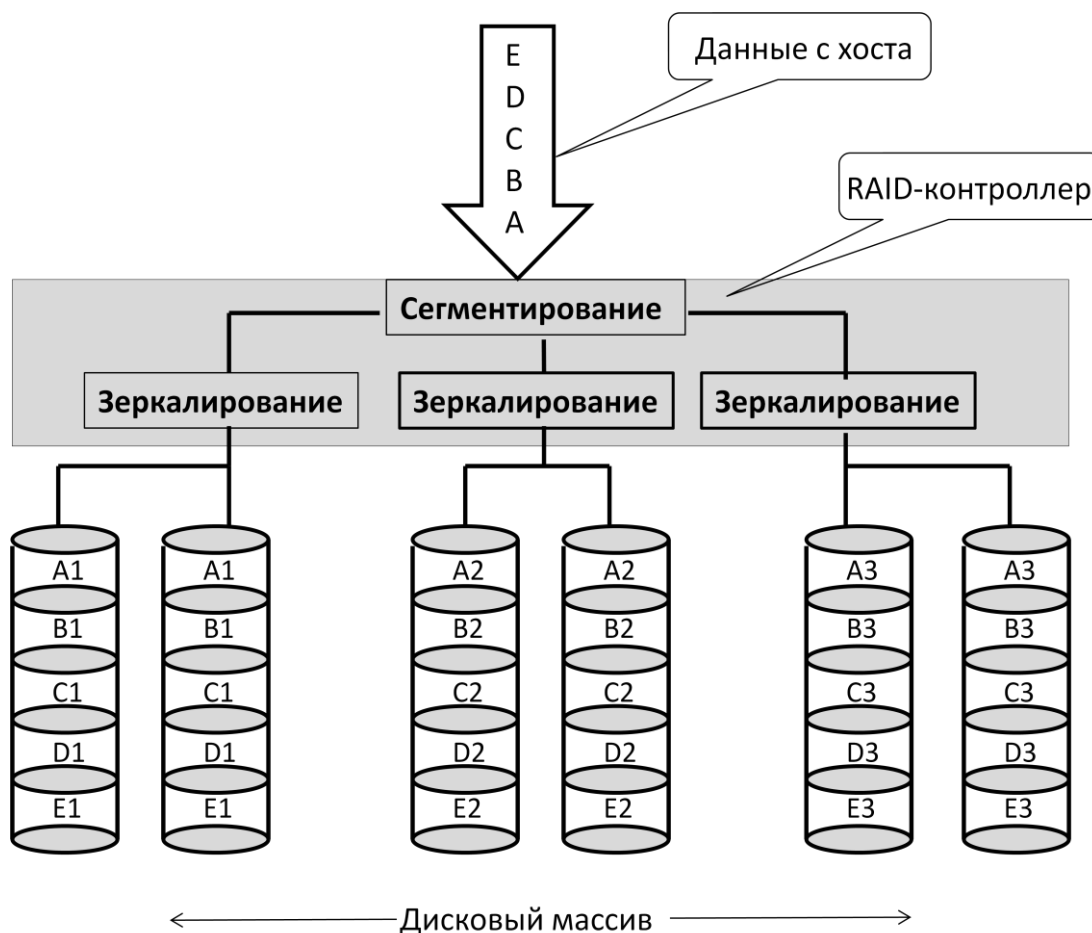


Рис. 4.8. Общая структура распределения данных при использовании RAID 01 (0+1)

Базовым элементом в RAID 01 является дорожка, то есть сначала данные сегментируются по разным дискам, а потом вся дорожка зеркалируется. При повреждении одного диска повреждается вся дорожка. Операция восстановления копирует всю дорожку на массив с поврежденным диском, что существенно увеличивает время восстановления данных.

Комбинированные уровни наследуют как преимущества, так и недостатки своих «родителей»: появление чередования в уровне RAID 50 несколько не добавляет ему надежности, но зато положительно отражается на производительности. Уровень RAID 15, будет очень надежным, но не самым быстрым и, к тому же, крайне неэкономичным: полезная ёмкость тома меньше половины суммарной ёмкости дисков.

#### 4.4. IOPS-операций и конфигурация дисков

При выборе конфигурации RAID-массива важно учитывать влияние структуры этого массива на производительность дискового массива. Так в RAID-массивах с зеркалированием, а также в массивах с контролем по четности на все операции записи необходимо тратить дополнительные ресурсы дисков, которые называются дополнительными расходами на запись или пенальти.

В RAID 1 все записи должны выполняться на два диска (основной и зеркало), таким образом, дополнительные расходы на запись составляют одну операцию.

В RAID 5 контроллер при выполнении каждой операции записи должен считывать, рассчитывать и записывать блок данных с информацией по четности. При наличии в RAID 5 четырех дисков, три диска используется для хранения данных, а четвертый – для хранения блока с контрольной информацией по четности (рис. 18). Информация  $A_p$  рассчитывается следующим образом

$$A_p = A_1 \oplus A_2 \oplus A_3$$

Если хост хочет изменить блок данных, который занимает место только на одном диске  $A_3$  (то есть на одном стрипе в пределах одного страйпа), то RAID контроллер не может просто записать этот небольшой блок данных на  $A_3$  и считать запрос хоста выполненным. Эта операция должна обновить данные четности  $A_p$ .

Новая информация рассчитывается следующим образом

$$A_{pnew} = A_{pold} - A_{3old} + A_{3new}$$

В последнем выражении должны выполняться операции XOR.

В булевой арифметике минус реализуется стандартной операцией исключающего ИЛИ, то есть

$$A_{pnew} = A_{pold} \oplus A_{3old} \oplus A_{3new}$$

После расчета новой контрольной информации по четности контроллер производит запись новых данных на  $A_3$  и  $A_p$ , на этом операция записи завершается (получается две операции записи).

Таким образом, для каждой операции записи блока данных, контроллер выполняет две операции считывания и две операции записи, дополнительные расходы на запись при использовании RAID 5 составляют 4 (рис.4.9).

Для определения количества дисков, требующихся для обеспечения работы приложения необходимо учитывать влияние RAID конфигурации на IOPS дисковой системы.

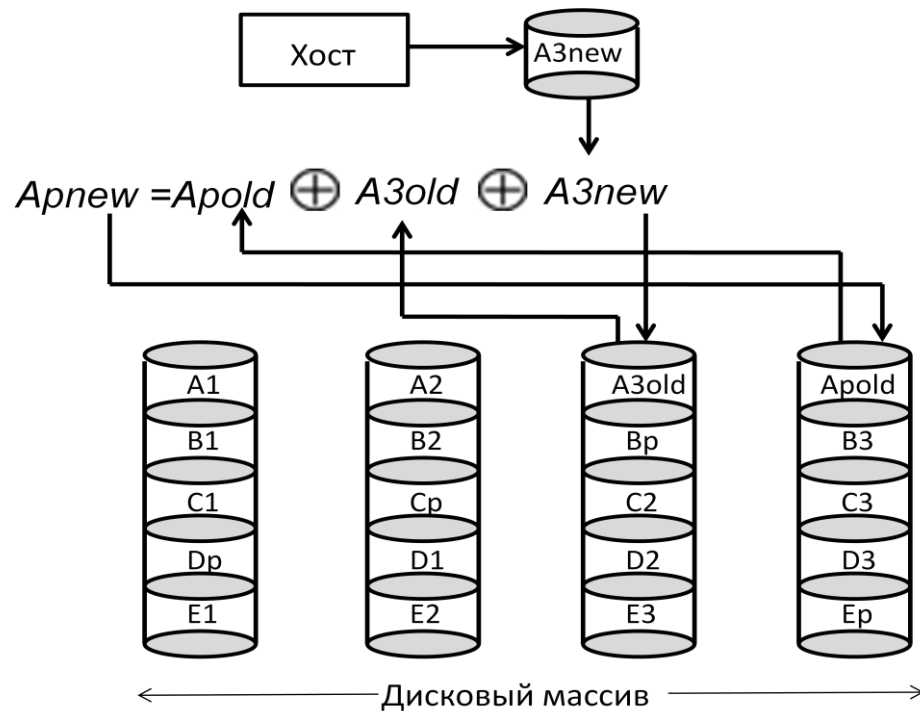


Рис. 4.9. Дополнительные расходы на запись в конфигурации RAID 5

Рассмотрим приложение, генерирующее 5200 IOPS операций, 60% из которых являются операциями считывания. Нагрузка на дисковую систему при наличии RAID 5 рассчитывается следующим образом.

$$RAID\ 5 = 0,6 \cdot 5200 + 4 \cdot (0,4 \cdot 5200) = 11\ 440\ IOPS$$

Рассчитанная нагрузка на дисковую систему позволяет определить количество дисков в системе.

Так если в данном примере используются жесткие диски с максимальной рабочей нагрузкой 180 IOPS, то количество дисков определится как

$$N_{hdd} = 11\ 400 / 180 = 64\ \text{диска}$$

Если использовать конфигурацию RAID 1, то нагрузка на дисковую систему составит

$$RAID\ 1 = 0,6 \cdot 5200 + 2 \cdot (0,4 \cdot 5200) = 7\ 200\ IOPS.$$

При использовании для реализации RAID 1 жестких дисков с такими же, как и в предыдущем случае, характеристиками их количество определится как



$$N_{hdd} = 7\,800 / 180 = 42 \text{ диска}$$

Здесь для получения результата количество дисков округляется до следующего четного числа.

#### **4.5. Влияние наличия RAID на производительность дисковых устройств**

В системах хранения данных RAID-массив является неотъемлемой составляющей, благодаря которому повышается пропускная способность и надежность данных, что необходимо для каждого предприятия ввиду ценности корпоративной информации. Также преимуществом является упрощение доступа к данным, что особенно необходимо для разделения ресурсов.

Одной из самых сложных проблем, с которыми сталкиваются организации в наши дни, является выбор и настройка действительно надежной сетевой системы хранения данных, а так же подобрать уровень RAID, обеспечивающий нужную производительность и надежность данных. Для достижения этого нужно определить цели, для которых будет служить RAID-массив, изучив тонкости каждого уровня, описанные выше.

Для измерения и оценки производительности жестких дисков, твердотельных дисков, составляющих RAID-массив, а также систем хранения данных используют такой параметр как **IOPS** (Input/Output Operations Per Second) – количество операций ввода/вывода в секунду. Перейдем к расчетам данной величины.

Основными измеряемыми величинами являются операции последовательного и произвольного доступа. Последовательный доступ – доступ к данным, осуществляемый в заранее заданном порядке. Произвольный доступ – доступ к данным, обращающийся к любому элементу последовательности за равные промежутки времени, не зависящие от размеров последовательности.

Для транзакционных систем со случайным доступом характерно большое количество запросов на операции чтения и записи небольшими блоками данных, например, база данных, имеющая размер блока 8-32 Кбайт.

Потоковый доступ – это последовательная запись или чтение блоков данных большого размера. Для потоковых приложений характерны запросы на операции с большими блоками данных (от 512 Кбайт и более).

Исходя из данной информации, можно выделить основные разновидности и характеристики IOPS:

- Random Read IOPS – число операций последовательного чтения в секунду;
- Sequential Read IOPS – число операций произвольного чтения в секунду;
- Sequential Write IOPS – число операций последовательной записи в секунду;

- Random Write IOPS – число операций произвольной записи в секунду;
- Total IOPS – суммарное число операций ввода/вывода в секунду.

Для того чтобы рассчитать производительность RAID-массива, изначально нужно знать величину IOPS отдельного диска.

Основными параметрами жесткого диска, задающими величину в IOPS, является RPM, время перехода магнитной головки (seek) и время задержки (latency). При помощи этих величин можно вычислить IOPS конкретного диска (так как IOPS вычисляется в секундах, каждая величина будет поделена на 1000):

$$Disk\ IOPS = RPM / ((seek\ time\ in\ ms / 1000) + (latency\ in\ ms / 1000)).$$

Для примера расчета был выбран жесткий диск Seagate SAS 15k RPM, модели ST3600957FC. По характеристикам от производителя, у данного диска время перехода магнитной головки на чтение/запись составляет 3.4/3.9 и время задержки – 2.0. Подставив все данные можно увидеть количество IOPS, которые может выдать данный жесткий диск.

$$Disk\ IOPS = 1.5 / (((Seek\ Time\ Average\ Read + Seek\ Time\ Average\ Write) / 2) / 1000) + (Average\ Latency / 1000) = 1.5 / ((3.65 / 1000) + (2 / 1000)) = 1.5 / ((0.00365) + (0.002)) = 265.48 \sim \underline{\underline{265\ IOPS}}$$

Так как жесткие диски обладают разным временем перехода магнитной головки и задержки, для расчетов будут использоваться следующие условные показатели (табл. 4.1):

Таблица 4.1

Условные показатели производительности накопителей

Скорость вращения диска	Интерфейс	IOPS
SSD	SATA 6 Гбит/с	55000
SAS 15K	SAS	265
SAS 10K	SAS	176
SATA 10K	SATA 6 Гбит/с	125
SATA 7.2K	SATA 6 Гбит/с	75

Зная IOPS жестких дисков, входящих в состав RAID-массива, можно переходить к расчетам самого массива.

Рассчитывая величину IOPS для определенного RAID-массива, следует учитывать его накладные расходы, а так же учет четности ввиду особенности каждого уровня.

Конечная формула выглядит следующим образом:

$$\text{Total IOPS} = R \text{ IOPS} + W \text{ IOPS}$$

$$R \text{ IOPS} = (\text{Disk IOPS} * (N - P_r)) * \%R$$

$$W \text{ IOPS} = ((\text{Disk IOPS} * (n - P_w)) * \%W) / C$$

Далее следует детально рассмотреть формулу:

*RIOPS* – количество IOPS при чтении;

*WIOPS* – количество IOPS при записи;

*DiskIOPS* – производительность в IOPS отдельного диска;

*N* – количество дисков в массиве;

*P<sub>r</sub>* – учет четности при чтении;

*P<sub>w</sub>* – учет четности при записи;

*C* – накладные расходы RAID-массива;

*%R* – общий процент чтения;

*%W* – общий процент записи;

*TotalIOPS* – суммарное количество IOPS, которое может выдать RAID-массив.

Параметры *P<sub>r</sub>* и *P<sub>w</sub>* являются уникальными для каждого уровня RAID, учитывающий четность. Данный параметр показывает накладные расходы определенного уровня, который вычитается из общего количества дисков и в результате расчет производится только с полезными дисками. Ниже предоставлена таблица накладных расходов для каждого уровня, использующего четность (табл.4.2):

Таблица 4.2

Накладные расходы для уровней, использующих четность

Уровень / Параметр	Pr	Pw
RAID 4	1	0
RAID 5	1	0
RAID 6	2	0

Следующий параметр, который следует рассмотреть – *C* (накладные расходы RAID-массива). В данном случае нужно учитывать расходы при записи на поддержание требуемого уровня массива. К примеру, чтобы записать данные в случае уровня RAID 5 происходит 4 операции:

1. чтение существующих данных;
2. четность RAID;
3. запись новых данных;
4. запись новой четности.

Ниже приведена табл. 4.3 значений накладных расходов при записи на каждый уровень RAID.

Таблица 4.3

Накладные расходы при записи данных на дисковые массивы

Уровень RAID	Накладные расходы
RAID 0	1
RAID 1	2
RAID 5	4
RAID 6	6
RAID 10	2

Для наглядности произведем расчет массива RAID 5 того же жесткого диска, который был взят в предыдущем примере.

#### Пример расчетов - RAID 5.

10 дисков. 15к. 70%W, 30%R.

Начальные данные, которые нам понадобятся – это IOPS данного диска (265), а так же все вышеописанные расходы на чтение и запись.

$$\text{Total IOPS} = R \text{ IOPS} + W \text{ IOPS}$$

$$R \text{ IOPS} = (\text{Disk IOPS} * (N - P_r)) * \%R$$

$$W \text{ IOPS} = ((\text{Disk IOPS} * (n - P_w)) * \%W) / C$$

$$R \text{ IOPS} = (265 * (10 - 1)) * 0.3 = 715.5 \text{ IOPS}$$

$$W \text{ IOPS} = (265 * (10 - 0)) * 0.7 / 4 = 463.75 \text{ IOPS}$$

$$\text{Total IOPS} = 1179.25 \sim 1180 \text{ IOPS}$$