

RAID

Материал из Википедии — свободной энциклопедии

Текущая версия страницы пока [не проверялась](#) опытными участниками и может значительно отличаться от [версии](#), проверенной 9 февраля 2013; проверки требуют [29 правок](#).

Перейти к: [навигация](#), [поиск](#)

У этого термина существуют и другие значения, см. [RAID \(значения\)](#).

RAID ([англ.](#) *redundant array of independent disks* — избыточный [массив независимых дисков](#)) — [массив](#) из нескольких дисков ([запоминающих устройств](#)), управляемых [контроллером](#), связанных между собой скоростными каналами передачи данных и воспринимаемых внешней системой как единое целое. В зависимости от типа используемого массива может обеспечивать различные степени отказоустойчивости и быстродействия. Служит для повышения надёжности хранения данных и/или для повышения скорости чтения/записи.

Аббревиатура «RAID» изначально расшифровывалась как «redundant array of inexpensive disks» («избыточный (резервный) массив недорогих дисков», так как они были гораздо дешевле RAM). Именно так был представлен RAID его создателями [Петтерсоном](#) (David A. Patterson), Гибсоном (Garth A. Gibson) и Катцом (Randy H. Katz) в 1987 году. Со временем «RAID» стали расшифровывать как «redundant array of independent disks» («избыточный (резервный) массив *независимых* дисков»), потому что для массивов приходилось использовать и дорогое оборудование (под недорогими дисками подразумевались диски для [ПЭВМ](#)).

[Калифорнийский университет в Беркли](#) представил ^{[источник не указан 28 дней](#)} следующие уровни спецификации RAID, которые были приняты как стандарт де-факто:

- **RAID 0** — дисковый массив повышенной производительности с чередованием, без отказоустойчивости;
- **RAID 1** — зеркальный дисковый массив;
- **RAID 2** зарезервирован для массивов, которые применяют [код Хемминга](#);
- **RAID 3 и 4** — дисковые массивы с чередованием и выделенным диском чётности;
- **RAID 5** — дисковый массив с чередованием и «невыделенным диском чётности»;
- **RAID 6** — дисковый массив с чередованием, использующий две контрольные суммы, вычисляемые двумя независимыми способами;
- **RAID 10** — массив RAID 0, построенный из массивов RAID 1;
- **RAID 50** — массив RAID 0, построенный из массивов RAID 5;
- **RAID 60** — массив RAID 0, построенный из массивов RAID 6.

Аппаратный RAID-контроллер может поддерживать несколько разных RAID-массивов одновременно, суммарное количество жёстких дисков которых не превышает количество разъёмов для них. При этом контроллер, встроенный в материнскую плату, в настройках BIOS имеет всего два состояния (включён или отключён), поэтому новый жёсткий диск, подключённый в незадействованный разъём контроллера при активированном режиме RAID, может игнорироваться системой, пока он не будет ассоциирован как ещё один RAID-массив типа JBOD (spanned), состоящий из одного диска.

Содержание

- [1 Базовые уровни RAID](#)
 - [1.1 RAID 0](#)
 - [1.2 RAID 1](#)
 - [1.3 RAID 2](#)
 - [1.4 RAID 3](#)
 - [1.5 RAID 4](#)
 - [1.6 RAID 5](#)
 - [1.7 RAID 6](#)
- [2 Комбинированные уровни](#)
 - [2.1 RAID 0+1](#)
 - [2.2 RAID 10](#)
- [3 Сравнение уровней RAID](#)
- [4 Нестандартные уровни RAID](#)
 - [4.1 RAID 7](#)
 - [4.2 RAID-DP](#)
- [5 Hybrid RAID](#)
- [6 Matrix RAID](#)
- [7 Дополнительные функции RAID-контроллеров](#)
- [8 Программный \(англ. *software*\) RAID](#)
- [9 Дальнейшее развитие идеи RAID](#)
- [10 Интересные факты](#)
- [11 См. также](#)
- [12 Примечания](#)
- [13 Ссылки](#)

Базовые уровни RAID

RAID 0

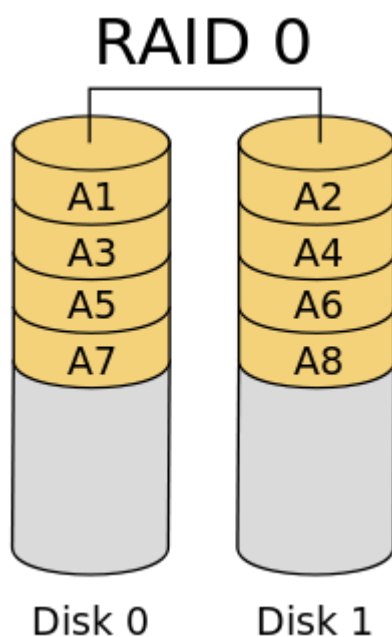




Схема RAID 0

RAID 0 (*striping* — «чередование») — дисковый массив из двух или более [жёстких дисков](#) без резервирования (т.е., по сути **RAID**-массивом не является). Информация разбивается на блоки данных (A_i) фиксированной длины и записывается на оба/несколько дисков одновременно.

(+): За счёт этого существенно повышается производительность (от количества дисков зависит кратность увеличения производительности).

(-): Надёжность RAID 0 заведомо ниже надёжности любого из дисков в отдельности и падает с увеличением количества входящих в RAID 0 дисков, т. к. отказ любого из дисков приводит к неработоспособности всего массива.

RAID 1



Два диска — минимальное количество для построения «зеркального» массива

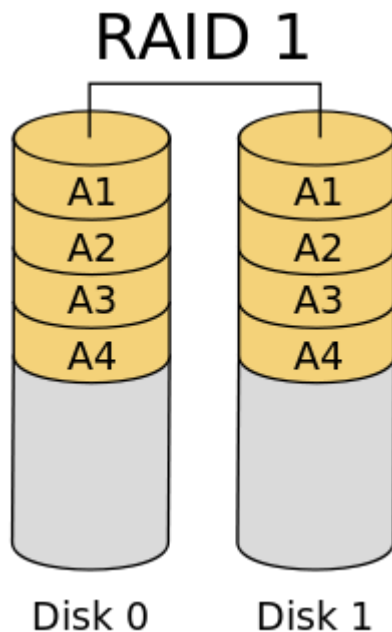


Схема RAID 1

RAID 1 (*mirroring* — «зеркалирование») — массив из двух дисков, являющихся полными копиями друг друга. Не следует путать с массивами RAID 1+0, RAID 0+1 и RAID 10, в которых используется более двух дисков и более сложные механизмы зеркалирования.

(+): Обеспечивает приемлемую скорость записи и выигрыш по скорости чтения при распараллеливании запросов. [\[1\]](#)

(+): Имеет высокую надёжность — работает до тех пор, пока функционирует хотя бы один диск в массиве. Вероятность выхода из строя сразу двух дисков равна [произведению вероятностей](#) отказа каждого диска, т.е. значительно ниже вероятности выхода из строя отдельного диска. На практике при выходе из строя одного из дисков следует срочно принимать меры — вновь восстанавливать избыточность. Для этого с любым уровнем RAID (кроме нулевого) рекомендуют использовать диски [горячего резерва](#).

(-): Недостаток RAID 1 в том, что по цене двух жестких дисков пользователь фактически получает лишь один.

RAID 2

Массивы такого типа основаны на использовании [кода Хемминга](#). Диски делятся на две группы: для данных и для кодов коррекции ошибок, причём если данные хранятся на $2^n - n - 1$ дисках, то для хранения кодов коррекции необходимо n дисков. Данные распределяются по дискам, предназначенным для хранения информации, так же, как и в RAID 0, т.е. они разбиваются на небольшие блоки по числу дисков. Оставшиеся диски хранят коды коррекции ошибок, по которым в случае выхода какого-либо жёсткого диска из строя возможно восстановление информации. Метод Хемминга давно применяется в памяти типа [ЕСС](#) и позволяет на лету исправлять однократные и обнаруживать двукратные ошибки.

Достоинством массива RAID 2 является повышение скорости дисковых операций по сравнению с производительностью одного диска.

Недостатком массива RAID 2 является то, что минимальное количество дисков, при котором имеет смысл его использовать, — 7. При этом нужна структура из почти двойного количества дисков (для $n=3$ данные будут храниться на 4 дисках), поэтому такой вид массива не получил распространения. Если же дисков около 30-60, то перерасход получается 11-19%.

RAID 3

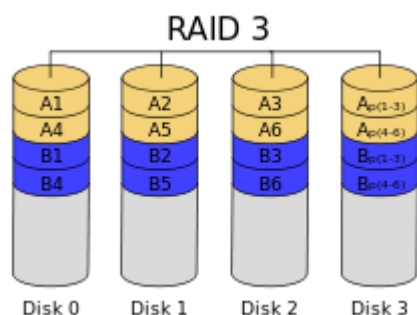


 Схема RAID 3

В массиве RAID 3 из n дисков данные разбиваются на куски размером меньше сектора (разбиваются на байты) или блоки и распределяются по $n - 1$ дискам. Ещё один диск используется для хранения блоков чётности. В RAID 2 для этой цели применялся $n - 1$ диск, но большая часть информации на контрольных дисках использовалась для коррекции ошибок на лету, в то время как большинство пользователей удовлетворяет простое восстановление информации в случае поломки диска, для чего хватает информации, уместяющейся на одном выделенном жёстком диске.

Отличия RAID 3 от RAID 2: невозможность коррекции ошибок на лету и меньшая избыточность.

Достоинства:

- высокая скорость чтения и записи данных;
- минимальное количество дисков для создания массива равно трём.

Недостатки:

- массив этого типа хорош только для однозадачной работы с большими файлами, так как время доступа к отдельному сектору, разбитому по дискам, равно максимальному из интервалов доступа к секторам каждого из дисков. Для блоков малого размера время доступа намного больше времени чтения.
- большая нагрузка на контрольный диск, и, как следствие, его надёжность сильно падает по сравнению с дисками, хранящими данные.

RAID 4

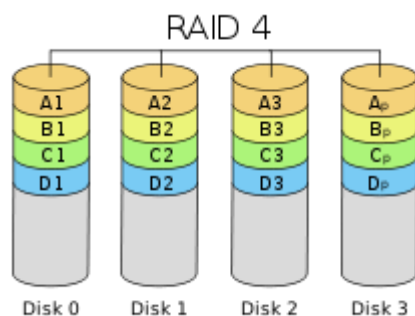


Схема RAID 4

RAID 4 похож на RAID 3, но отличается от него тем, что данные разбиваются на блоки, а не на байты. Таким образом, удалось отчасти «победить» проблему низкой скорости передачи данных небольшого объёма. Запись же производится медленно из-за того, что чётность для блока генерируется при записи и записывается на единственный диск. Из систем хранения широкого распространения RAID-4 применяется на устройствах хранения компании [NetApp](#) (NetApp FAS), где его недостатки успешно устранены за счет работы дисков в специальном режиме групповой записи, определяемом используемой на устройствах внутренней файловой системой [WAFL](#).

RAID 5

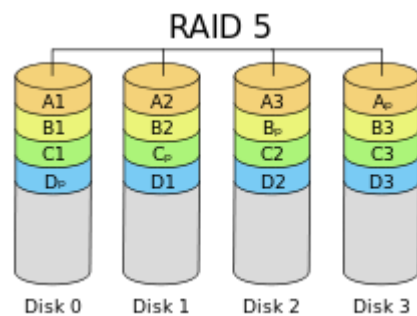


Схема RAID 5

Основным недостатком уровней RAID от 2-го до 4-го является невозможность производить параллельные операции записи, так как для хранения информации о чётности используется отдельный контрольный диск. RAID 5 не имеет этого недостатка. Блоки данных и контрольные суммы циклически записываются на все диски массива, нет асимметрии конфигурации дисков. Под контрольными суммами подразумевается результат операции [XOR](#) (исключающее или). *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, мы можем получить его, 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 хог требуется всего 1 диск, размер которого равен размеру любого другого диска в raid.

(+): RAID5 получил широкое распространение, в первую очередь, благодаря своей экономичности. Объём дискового массива RAID5 рассчитывается по формуле $(n-1) \cdot \text{hddsize}$, где n — число дисков в массиве, а hddsize — размер наименьшего диска. Например, для массива из 4-х дисков по 80 гигабайт общий объём будет $(4 - 1) \cdot 80 = 240$ гигабайт. На запись информации на том RAID 5 тратятся дополнительные ресурсы и падает производительность, так как требуются дополнительные вычисления и операции записи, зато при чтении (по сравнению с отдельным винчестером) имеется выигрыш, потому что потоки данных с нескольких дисков массива могут обрабатываться параллельно.

(-): Производительность RAID 5 заметно ниже, в особенности на операциях типа Random Write (записи в произвольном порядке), при которых производительность падает на 10-25% от производительности RAID 0 (или RAID 10), так как требует большего количества операций с дисками (каждая операция записи сервера заменяется на контроллере RAID на три - одну операцию чтения и две операции записи). Недостатки RAID 5 проявляются при выходе из строя одного из дисков — весь том переходит в критический режим (degrade), все операции записи и чтения сопровождаются дополнительными манипуляциями, резко падает производительность. При этом уровень надёжности снижается до надёжности RAID-0 с соответствующим количеством дисков (то есть в n раз ниже надёжности одиночного диска). Если до полного восстановления массива произойдет выход из строя, или возникнет невозможная ошибка чтения хотя бы на еще одном диске, то массив разрушается, и данные на нем восстановлению обычными методами не подлежат. Следует также 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 повышает риск сбоя при восстановлении данных.

Минимальное количество используемых дисков равно трём.

RAID 6

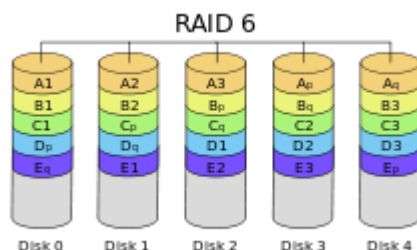


Схема RAID 6

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. Требуется более мощный RAID-контроллер. Обеспечивает работоспособность после одновременного выхода из строя двух дисков — защита от кратного отказа. Для организации массива требуется минимум 4 диска^[2]. Обычно использование RAID-6 вызывает примерно 10-15% падение производительности дисковой группы, по сравнению с аналогичными показателями RAID-5, что вызвано большим объёмом обработки для контроллера (необходимость рассчитывать вторую контрольную сумму, а также прочитывать и перезаписывать больше дисковых блоков при записи каждого блока).

Комбинированные уровни

Помимо базовых уровней RAID 0 - RAID 6, описанных в [стандарте «Common RAID Disk Drive Format \(DEF\) standard»](#), существуют комбинированные уровни с названиями вида «RAID $\alpha+\beta$ » или «RAID $\alpha\beta$ », что обычно означает «RAID β , составленный из нескольких RAID α » (иногда производители интерпретируют это по-своему).

Например:

- RAID 10 (или 1+0) — это RAID 0, составленный из нескольких (или хотя бы двух) RAID 1 (зеркалированных пар).
- RAID 51 — RAID 1, зеркалирующий два RAID 5.

Комбинированные уровни наследуют как преимущества, так и недостатки своих «родителей»: появление чередования в уровне RAID 5+0 несколько не добавляет ему надёжности, но зато положительно отражается на производительности. Уровень RAID 1+5, наверное, очень надёжный, но не самый быстрый и, к тому же, крайне неэкономичный: полезная ёмкость тома меньше половины суммарной ёмкости дисков...

RAID 0+1

Под RAID 0+1 может подразумеваться два варианта:

- два RAID 0 объединяются в RAID 1;
- в массив объединяются три и более диска, и каждый блок данных записывается на два диска данного массива^[3]; таким образом, при таком подходе, как и в «чистом» RAID 1, полезный объём массива составляет половину от суммарного объёма всех дисков (если это диски одинаковой ёмкости).

RAID 10

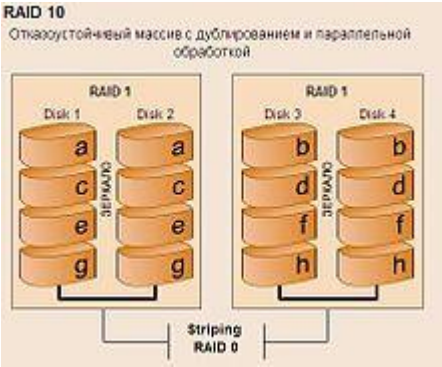


Схема архитектуры RAID 10

RAID 10 — зеркалированный массив, данные в котором записываются последовательно на несколько дисков, как в [RAID 0](#). Эта архитектура представляет собой массив типа RAID 0, сегментами которого вместо отдельных дисков являются массивы RAID 1. Соответственно, массив этого уровня должен содержать как минимум 4 диска (и всегда чётное количество). RAID 10 объединяет в себе высокую отказоустойчивость и производительность.

Утверждение, что RAID 10 является самым надёжным вариантом для хранения данных, ошибочно, т.к., несмотря на то, что для данного уровня RAID возможно сохранение целостности данных при выходе из строя половины дисков, необратимое разрушение массива происходит при выходе из строя уже двух дисков, если они находятся в одной зеркальной паре.

Сравнение уровней RAID

Уровень	Количество дисков	Эффективная ёмкость*	Отказоустойчивость	Преимущества	Недостатки
0	от 2	$S * N$	нет	наивысшая производительность	очень низкая надёжность
1	2	S	1 диск	надёжность	двойная стоимость дискового пространства
1E	от 3	$S * N / 2$	1 диск**	высокая защищённость данных и неплохая производительность	двойная стоимость дискового пространства
10	от 4, чётное	$S * N / 2$	1 диск***	наивысшая производительность и высокая надёжность	двойная стоимость дискового пространства
5	от 3 до 16	$S * (N - 1)$	1 диск	экономичность, высокая надёжность, неплохая производительность	производительность ниже RAID 0

50	от 6, чётное	$S * (N - 2)$	1 диск***	высокая надёжность и производительность	высокая стоимость и сложность обслуживания
5E	от 4	$S * (N - 2)$	1 диск	экономичность, высокая надёжность, скорость выше RAID 5	производительность ниже RAID 0 и 1, резервный накопитель работает на холостом ходу и не проверяется
5EE	от 4	$S * (N - 2)$	1 диск	быстрое реконструирование данных после сбоя, экономичность, высокая надёжность, скорость выше RAID 5	производительность ниже RAID 0 и 1, резервный накопитель работает на холостом ходу и не проверяется
6	от 4	$S * (N - 2)$	2 диска	экономичность, наивысшая надёжность	производительность ниже RAID 5
60	от 8, чётное	$S * (N - 2)$	2 диска	высокая надёжность, большой объем данных	высокая стоимость и сложность организации
61	от 8, чётное	$S * (N - 2) / 2$	4 диска**	очень высокая надёжность	высокая стоимость и сложность организации

* N — количество дисков в массиве, S — объём наименьшего диска. [\[4\]\[5\]\[6\]\[7\]](#) ** Информация не потеряется, если выйдут из строя все диски в пределах одного зеркала. *** Информация не потеряется, если выйдут из строя два диска в пределах разных зеркал.

Нестандартные уровни RAID

RAID 7

RAID 7 - зарегистрированная торговая марка компании [Storage Computer Corporation](#), отдельным уровнем RAID не является. Структура массива такова: на $n - 1$ дисках хранятся данные, один диск используется для складирования блоков чётности. Запись на диски кешируется с использованием оперативной памяти, сам массив требует обязательного [ИБП](#); в случае перебоев с питанием происходит повреждение данных.

RAID-DP

Существует модификация RAID-6 компании NetApp - RAID-DP. Отличие от традиционного массива заключается в выделении под контрольные суммы двух отдельных дисков. Благодаря взаимодействию RAID-DP и файловой системы [WAFL](#) (все операции записи последовательны и производятся на свободное место) пропадает падение производительности как в сравнении с RAID-5, так и в сравнении с RAID-6.

Hybrid RAID

«Hybrid RAID» — это некоторые из обычных уровней RAID, но в сочетании с дополнительным ПО и SSD-дисками, которые используются как кэш для чтения. В результате производительность системы повышается, т.к. SSD-диски обладают значительно лучшими скоростными характеристиками по сравнению с HDD. Существует

несколько реализаций, например Crucial Adrenaline, либо некоторые контроллеры Adaptec бюджетного класса. На данный момент Hybrid RAID не рекомендуется использовать в серверах ввиду малого ресурса SSD-дисков.

Matrix RAID

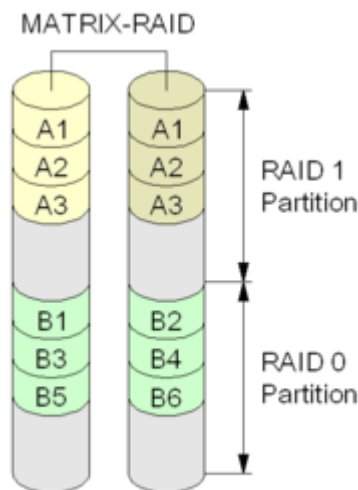


Схема [Intel](#) Matrix RAID

Matrix RAID — это технология, реализованная фирмой [Intel](#) в своих [чипсетах](#), начиная с [ICH6R](#). Эта технология не является новым уровнем RAID (и ее аналог существует в аппаратных RAID-контроллерах высокого уровня), она позволяет, используя небольшое количество дисков, организовать на разных разделах этих дисков одновременно несколько массивов уровня RAID 1, RAID 0 и RAID 5. Это позволяет за сравнительно небольшие деньги обеспечить для одних данных повышенную надёжность, а для других — высокую производительность.

Дополнительные функции RAID-контроллеров

Многие RAID-контроллеры оснащены набором дополнительных функций:

- "Горячая замена" ([Hot Swap](#))
- "Горячий резерв" (Hot Spare)
- Проверка на стабильность.

Программный ([англ.](#) *software*) RAID

Для реализации RAID можно применять не только аппаратные средства, но и полностью программные компоненты ([драйверы](#)). Например, в системах на [ядре Linux](#) существуют специальные модули [ядра](#), а управлять RAID-устройствами можно с помощью утилиты [mdadm](#). Программный RAID имеет свои достоинства и недостатки. С одной стороны, он ничего не стоит (в отличие от аппаратных [RAID-контроллеров](#), цена которых от \$250). С другой стороны, программный RAID использует ресурсы [центрального процессора](#), и в моменты пиковой нагрузки на дисковую систему процессор может значительную часть мощности тратить на обслуживание RAID-устройств.

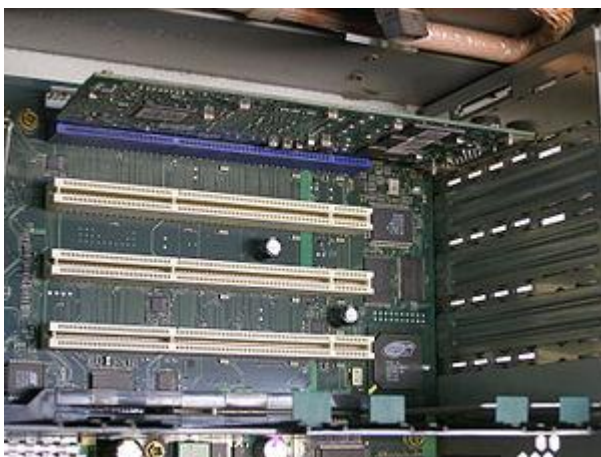
Ядро Linux 2.6.28 (последнее из вышедших в 2008 году) поддерживает программные RAID следующих уровней: 0, 1, 4, 5, 6, 10. Реализация позволяет создавать RAID на отдельных разделах дисков, что аналогично описанному выше Matrix RAID. Поддерживается загрузка с RAID.

ОС семейства [Windows NT](#), такие как Windows NT 3.1/3.5/3.51/[NT4/2000/XP/2003](#) изначально, с момента проектирования данного семейства, поддерживает программный RAID 0, RAID 1 и RAID 5 (см. [Dynamic Disk](#)). Более точно, Windows XP Pro поддерживает RAID 0. Поддержка RAID 1 и RAID 5 заблокирована разработчиками, но, тем не менее, может быть включена, путем редактирования системных бинарных файлов ОС, что запрещено лицензионным соглашением.^[8] Windows 7 поддерживает программный RAID 0 и RAID 1, Windows Server 2003 — 0, 1 и 5. Windows XP Home не поддерживает RAID.

В ОС [FreeBSD](#) есть несколько реализаций программного RAID. Так, `atacontrol`, может как полностью строить программный RAID, так и может поддерживать полуаппаратный RAID на таких чипах как ICH5R. Во FreeBSD, начиная с версии 5.0, дисковая подсистема управляется встроенным в ядро механизмом GEOM. GEOM предоставляет модульную дисковую структуру, благодаря которой родились такие модули как `gstripe` (RAID 0), `gmirror` (RAID 1), `graid3` (RAID 3), `gconcat` (объединение нескольких дисков в единый дисковый раздел). Также существуют устаревшие классы `ccd` (RAID 0, RAID 1) и `gvinum` (менеджер логических томов `vinum`). Начиная с FreeBSD 7.2 поддерживается файловая система [ZFS](#), в которой можно собирать следующие уровни RAID: 0, 1, 5, 6, а также комбинируемые уровни.

[OpenSolaris](#) и [Solaris 10](#) используют [Solaris Volume Manager](#), который поддерживает RAID-0, RAID-1, RAID-5 и любые их комбинации как 1+0. Поддержка RAID-6 осуществляется в файловой системе [ZFS](#).

Дальнейшее развитие идеи RAID



Синий разъём PCI-X на материнской плате сервера FSC Primergy TX200 S2 специально предназначен для платы ноль-канального RAID (zero-channel RAID, ZCR). Установлен MegaRAID [320-0](#) Zero Channel RAID Controller фирмы LSI^[9]

Идея RAID-массивов — в объединении дисков, каждый из которых рассматривается как набор секторов, и в результате драйвер файловой системы «видит» как бы единый диск и работает с ним, не обращая внимания на его внутреннюю структуру. Однако, можно добиться существенного повышения производительности и надёжности дисковой

системы, если драйвер файловой системы будет «знать» о том, что работает не с одним диском, а с набором дисков.

Более того: при разрушении любого из дисков в составе RAID-0 вся информация в массиве окажется потерянной. Но если драйвер файловой системы разместил каждый файл на одном диске, и при этом правильно организована структура директорий, то при разрушении любого из дисков будут потеряны только файлы, находившиеся на этом диске; а файлы, целиком находящиеся на сохранившихся дисках, останутся доступными.

Интересные факты

- Сотрудник корпорации Y-E Data, которая является крупнейшим в мире производителем USB флoppi-дисководов, Дэниэл Олсон в качестве эксперимента создал RAID-массив из четырёх [iPod Shuffle](#)^[10].
- Годовая вероятность отказа дискового массива RAID 5 из 3 дисков [WD Caviar Blue](#) = 0,05 %^[11].

См. также

- [JBOD](#) — дисковый массив с последовательным распределением дискового пространства по дискам.
- [NAS](#) — внешнее сетевое хранилище данных со встроенным RAID.

Примечания

- ↑ [R.LAB. RAID-массивы — надёжность и производительность. RAID 1](#)
- ↑ [RAID Level 6](#) (англ.)(недоступная ссылка — *история*). pcguide.com. Проверено 7 июля 2011.
- ↑ [Выбираем уровень RAID](#)
- ↑ [RAID 0, RAID 1, RAID 5, RAID6, RAID 10 или что такое уровни RAID?](#)
- ↑ [Dynamic disk introduction - RAID levels](#)
- ↑ [RAID Array and Server: Hardware and Service Comparison](#)
- ↑ [Summary Comparison of RAID Levels](#)
- ↑ [Создаём программный массив RAID 5 под Windows XP](#)
- ↑ [LSI MegaRAID 320-0 Zero Channel RAID Controler User's Guide](#) (англ.)
- ↑ [iPod RAID](#)
- ↑ [Формула надёжности RAID](#)

Ссылки



- [Новые уровни RAID: цифры, буквы и то, что за ними](#)
- [Неизбыточно о RAID \(XOR, RAID-Z\)](#)
- [All RAID Levels Description by iXBT.com](#)

- [Пример восстановления данных \(виртуально\) с неисправного NAS Raid массива 1Tb](#)
- [Пример восстановления данных с Raid 0 при помощи WinHex](#)