

Организация RAID массивов

Неотъемлемую роль в нашем компьютере играют жесткие диски, ведь на них хранится вся информация, с них запускается операционная система, в них обитает файл подкачки и прочее, прочее, прочее, прочее.

Как известно, эти самые жесткие диски так же имеют некий запас прочности после которого выходят из строя, а так же характеристики влияющие на производительность.



Что такое RAID и зачем оно нужно.

RAID — это дисковый массив (т.е. комплекс или, если хотите, связка) из нескольких устройств, – жестких дисков.

Как я и говорил выше, этот массив служит для повышения надёжности хранения данных и/или для повышения скорости чтения/записи информации.

Собственно, то чем именно занимается оная связка из дисков, – ускорением работы или повышением безопасности данных, зависит от Вас, а точнее, от выбора текущей конфигурации рейда(ов). Разные типы этих конфигураций как раз и отмечаются разными номерами, – **1, 2, 3, 4** и пр, – и выполняют разные функции.

Рейды ощутимо удобнее и эффективнее использования одного диска в системе.

Я рекомендую их использование всем поголовно, не смотря на то, что приходится использовать два (а то и все четыре) устройства вместо одного.

В случае построения 0-вой версии (описание вариаций **0, 1, 2, 3** и пр., изучим далее) Вы получите **ощутимый** прирост производительности. Да и вообще жесткий диск нынче как раз таки узкий канал в быстродействии системы. Почему? Судите сами, – процессоры обзаводятся ядрами, частотами, кэшем и архитектурой; видеокарты, – числом пиксельных конвейеров, количеством и разрядностью памяти, шейдерными блоками, частотами видеопроцессоров и кое-где даже количеством этих процессоров; оперативная память, – частотами и [таймингами](#) ; жесткие диски же растут разве что в объеме ибо скорость оборота головки оных (за исключением редких моделей типа Raptor'ов) замерла уже довольно давно на отметке в **7200**, кэш тоже не то чтобы растет, архитектура остается почти прежней.. В общем в плане производительности диски стоят на месте (ситуацию могут спасти разве что развивающиеся SSD), а ведь они играют весомую роль в работе системы и, местами, полновесных приложений.

В случае же построения единичного (в смысле за номером **1**) рейда Вы чуток потеряете в производительности, но зато получите почти **100%**-ую гарантию безопасности Ваших данных, ибо оные будут до байтика полностью дублироваться и даже в случае выхода из строя одного диска, – все целиком и полностью будет находится на втором без всяких потерь. В общем, повторяюсь, рейды будут полезны всем и каждому. Я бы даже сказал, – обязательны.

Физически **RAID**-массив представляет собой от **двух** до **n**-го количества жестких дисков подключенных к мат.плате поддерживающей возможность создания **RAID** (или к соответствующему контроллеру, что реже ибо оные дороги для рядового пользователя (контроллеры обычно используются на серверах в силу повышенной надежности и производительности)), т.е. на глаз ничего внутри системника не изменяется, никаких лишних подключений или соединений дисков между собой или с чем-то еще попросту нет. В общем в аппаратной части всё почти как всегда, а изменяется лишь программный подход, который, собственно, и задает, путем выбора типа рейда, как именно должны работать подключенные диски.

Программно, в системе, после создания рейда, тоже не появляется никаких особенных причуд. По сути, вся разница в работе с рейдом заключается только в небольшой настройке в биосе, которая собственно организует рейд (см.ниже) и в использовании драйвера. В остальном ВСЁ совершенно тоже самое – в “Мой компьютер” те же **C**, **D** и прочие диски, всё те же папки, файлы.. В общем и программно, на глаз, полная идентичность.

Установка массива не представляет собой ничего сложного: просто берем мат.плату, которая поддерживает технологию **RAID**, берем два полностью идентичных, – **это важно!**, – как по характеристикам (размеру, кэшу, интерфейсу и пр) так и по производителю и модели, диска и подключаем их к одной мат.плате. Далее просто включаем компьютер, заходим в **BIOS** и выставляем параметр **SATA Configuration: RAID**.

После этого в процессе загрузки компьютера (как правило, до загрузки Windows) появляется панель отображающая информацию о дисках в рейде и вне него, где, собственно нужно нажать **CTR-I**, чтобы настроить рейд (добавить диски в него, удалить и тд и тп). Собственно, вот и все. Дальше идет установка **Windows** и прочие радости жизни, т.е, опять же, всё как всегда.

Важно: при создании или удалении рейда (1-го рейда это вроде не касается, но не факт) неизбежно удаляется вся информация с дисков, а посему просто проводить эксперимент, создавая и удаляя различные конфигурации, явно не стоит. Посему, перед созданием рейда предварительно сохраните всю нужную информацию (если она есть), а потом уже экспериментируйте.

Что до конфигураций.. Как я уже говорил, **RAID** массивов существует несколько видов (как минимум из основного базиса, – это RAID 1, RAID 2, RAID 3, RAID 4, RAID 5, RAID 6). Для начала я расскажу о двух, наиболее понятных и популярных среди обычных пользователей:

RAID 0 – дисковый массив для увеличения скорости\записи.

RAID 1 – зеркальный дисковый массив.

RAID 0

И так.. **RAID 0** (он же, страйп («Striping»)) — используется от двух до четырех (больше, – реже) жестких дисков, которые совместно обрабатывают информацию, что повышает производительность. Чтобы было понятно, – таскать мешки одному человеку дольше и сложнее чем вчетвером (хотя мешки остаются все теми же по своим физ свойствам, меняются лишь мощности с ними взаимодействующие). Программно же, информация на рейде такого типа, разбивается на блоки данных и записывается на оба/несколько дисков поочередно. Один блок данных на один диск, другой блок данных на другой и тд. Таким образом **существенно** повышается производительность (от количества дисков зависит кратность увеличения производительности, т.е 4-ые диска будут бегать шустрее чем два), но страдает безопасность данных на всём массиве. При выходе из строя любого из входящих в такой **RAID** винчестеров (т.е. жестких дисков) полностью и безвозвратно пропадает вся информация. Почему? Дело в том, что каждый файл состоит из некоторого количества байт.. каждый из которых несет в себе информацию. Но в **RAID 0** массиве байты одного файла могут быть расположены на нескольких дисках. Соответственно при “смерти” одного из дисков потеряется произвольное количество байтов файла и восстановить его будет просто невозможно. Но файл то не один.. В общем при использовании такого рейд-массива настоятельно рекомендуется делать постоянные [бэкапы](#) ценной информации на внешний носитель. Рейд действительно обеспечивает ощутимую скорость – это я Вам говорю на собственном опыте, т.к у меня дома уже годами установлено такое счастье.

RAID 1

Что же до RAID 1 (Mirroring — «зеркало»).. Собственно, начну с недостатка. В отличии от **RAID 0** получается, что Вы как бы “теряете” объем второго жесткого диска (он используется для записи на него полной (байт в байт) копии первого жесткого диска в то время как RAID 0 это место полностью доступно).

Преимущество же, как Вы уже поняли, в том, что он имеет высокую надежность, т.е. все работает (и все данные существуют в природе, а не исчезают с выходом из строя одного из устройств) до тех пор пока функционирует хотя бы один диск, т.е. если даже грубо вывести из строя один диск – Вы не потеряете ни байта информации, т.к. второй является чистой копией первого и заменяет его при выходе из строя. Такой рейд частенько используется в серверах в силу безумнейшей жизнеспособности данных, что важно.

При подобном подходе в жертву приносится производительность и, по личным ощущениям, она даже меньше чем при использовании одного диска без всяких там рейдов. Впрочем, для некоторых надежность куда важнее производительности.

RAID 2.

Описание этих массивов тут по столько по сколько, т.е. чисто для справки, да и то в сжатом (по сути описан только второй) виде. Почему так? Как минимум в силу низкой популярности этих массивов среди рядового (да и в общем-то любого другого) пользователя и, как следствие, малого опыта использования оных мною.

RAID 2 зарезервирован для массивов, которые применяют некий код Хемминга.

Принцип работы примерно такой: данные записываются на соответствующие устройства так же, как и в **RAID 0**, т.е. они разбиваются на небольшие блоки по всем дискам, которые участвуют в хранении информации. Оставшиеся же (специально выделенные под оное) диски хранят коды коррекции ошибок, по которым в случае выхода какого-либо винчестера из строя возможно восстановление информации. В массивах такого типа диски делятся на две группы — для данных и для кодов коррекции ошибок. Например, у Вас два диска являются собой место под систему и файлы, а еще два будут полностью отведены под данные коррекции на случай выхода из строя первых двух дисков. По сути это что-то вроде нулевого рейда, только с возможностью хоть как-то спасти информацию в случае сбоя одного из винчестеров. Редкостно затратно, — четыре диска вместо двух с весьма спорным приростом безопасности.

RAID3

Третий уровень использует чередование и выделенный диск для контроля четности. Блоки данных обычно имеют длину меньше 1024 байт.

Информация распределяется на несколько дисков, а высчитанное значение по четности сохраняется на отдельный диск. Все скоростные преимущества чередования сводятся к нулю из-за необходимости записывать контрольную сумму на выделенный диск, а больше всех страдает скорость случайной записи. К достоинствам отнесем возможность работы массива при отказе одного из дисков.

RAID4

Отличается от RAID 3 только размером блока данных при чередовании. Это несколько улучшает работу массива при случайном чтении, но запись все равно довольно медленная. Диск с контрольными суммами является ярко выраженным «узким местом» в системе. RAID4 не нашел своего места на рынке и редко используется. Это держит цены на соответствующие контроллеры на высоком уровне.

RAID5

Наиболее распространенный в системах хранения данных – пятый уровень. Он характеризуется применением чередования и четности. В отличие от RAID 3, контрольные суммы не хранятся на одном диске, а разбрасываются по всем, что позволяет значительно поднять скорость записи. Главный принцип распределения экстраблоков: они не должны располагаться на том же диске, с которого была зашифрована информация. Надежность и скорость работы такой системы оказываются очень даже высокими. При восстановлении информации всю работу на себя берет RAID контроллер, так что операция проходит довольно быстро.

А это, пожалуй, самый популярный вид RAID-массива, в связи с экономичностью использования носителей данных. Блоки данных и проверочные суммы циклически пишутся на все диски массива. Если один диск выходит из строя, конечно снизится производительность, но данные не пропадут.

Под проверочными суммами подразумевается результат операции XOR.

$\text{Check 1} = \text{Data 1 XOR Data 2 XOR Data 3 XOR Data 4}.$

Теперь, если вдруг у нас пропадет 3-ий диск (бдыщ! Взорвался). То получить

Data 3 можно так:

$\text{Data 3} = \text{Data 1 XOR Data 2 XOR Check 1 XOR Data 4}.$

RAID6

Для некоторых особо критичных приложений требуется повышенная надежность. Например, чтобы при выходе из строя даже двух дисков массив сохранил данные и даже остался работоспособным. Можно ли это сделать? Конечно, решение лежит на поверхности. используются все те же технологии чередования и четности. Но контрольная сумма вычисляется два раза и копируется на два разных диска. В итоге данные окажутся потерянными только в случае выхода из строя сразу трех жестких дисков .

По сравнению с RAID 5 это более дорогое и медленное решение, которое может показать себя разве что при случайном чтении. На практике RAID 6 почти не используется, так как выход из строя сразу двух дисков – слишком редкий случай, а повысить надежность можно другими способами.

RAID7

В отличие от остальных уровней, RAID 7 не является открытым стандартом, столь звучное и выгодное название выбрала для своей модификации RAID 3 компания Storage Computer Corporation. Улучшения заключаются в использовании асинхронного чередования, применении кэш-памяти и специального высокопроизводительного микропроцессора. Обеспечивая такой же, как в RAID 3, уровень надежности, RAID 7 значительно выигрывает в скорости. Недостаток у него один, но очень серьезный – огромная цена, обусловленная монополией на изготовление контроллеров.

Комбинированные RAID-массивы

RAID 0+1 и 10

RAID 0+1 часто называют «зеркалом чередований», а RAID 10 – «чередованием зеркал». В обоих случаях используются две технологии – чередование и зеркалирование, но результаты разные. RAID 0+1 обладает высокой скоростью работы и повышенной надежностью, поддерживается даже дешевыми RAID контроллерами и является недорогим решением. Но по надежности несколько лучше RAID 10. Так, массив из 10 дисков (5 по 2) может остаться работоспособным при отказе до 5 жестких дисков! Основным недостатком этих массивов – низкий процент использования емкости накопителей – всего 50%. Но для домашних систем именно RAID 01 или 10 может оказаться оптимальным решением.

RAID 0+3 и 30

С этими массивами у производителей наблюдается путаница. Довольно часто вместо 0+3 или 30 указывают более привлекательное число 5+3 (53). Не верьте! По идее сочетание чередования и RAID 3 дает выигрыш в скорости, но он довольно мал. Зато система заметно усложняется. Наиболее простой уровень 30. Из двух массивов RAID 3 строится страйп (чередование), и минимальное количество требуемых дисков – 6. Получившийся RAID 30 с точки зрения надежности лучше, чем 0+3. Достоинства этих комбинаций в довольно высоком проценте использования емкости дисков и высокой скорости чтения данных. Недостатки – высокая цена, сложность системы.

RAID 0+5 и 50

Что будет, если объединить чередование с распределенной четностью с обыкновенным чередованием? Получится быстрая и надежная система. RAID 0+5 представляет собой набор страйпов, на основе которых построен RAID 5. Такая комбинация используется редко, так как практически не дает выигрыша ни в чем. Широкое распространение получил составной RAID массив 50. Чаще всего это два массива RAID 5, объединенных в страйп. Такая конфигурация позволяет получить высокую производительность при работе с файлами малого размера. Типичный пример – использование в качестве WEB-сервера.

RAID 1+5 (15) и 5+1 (51)

Этот уровень построен на сочетании зеркалирования или дуплекса и чередования с распределенной четностью. Основная цель RAID 15 и 51 – значительное повышение надежности. Массив 1+5 продолжает работать при отказе трех накопителей, а 5+1 – даже при потере пяти из восьми жестких дисков! Платить приходится большим количеством неиспользуемой емкости дисков и общим удорожанием системы. Чаще всего для построения RAID 5+1 используют два контроллера RAID 5, которые зеркалируют на программном уровне, что позволяет снизить затраты.

Matrix RAID

Matrix RAID — это технология, реализованная фирмой Intel в своих чипсетах начиная с ICH6R. Строго говоря, эта технология не является новым уровнем RAID (ее аналог существует в аппаратных RAID-контроллерах высокого уровня), она позволяет, используя лишь 2 диска, организовать одновременно один или несколько массивов уровня RAID 1 и один или несколько массивов уровня RAID 0. Это позволяет за сравнительно небольшие деньги обеспечить для одних данных повышенную надёжность, а для других высокую скорость доступа.

JBOD

А что делать, если нужен просто один логический диск гигантского размера? Без всяких зеркалирований, чередования и четности? Тогда это уже не RAID, а JBOD – Just A Bunch Of Disks. Реализовать этот режим способен простейший контроллер или даже программная реализация контроллера. Есть ли у него преимущества, если JBOD не повышает ни быстродействия, ни надёжности? Есть. По крайней мере, для работы используется все доступное пространство жестких дисков. И еще: в случае выхода из строя одного из жестких дисков, информация на других не повреждается.

	RAID0	RAID1	RAID3	RAID5	RAID6	RAID0+1	RAID10	RAID50	RAID51
Минимальное количество жестких дисков	2,4	2		3			4		6
Доступное рабочее пространство, %	100	50	66 для 3х, 75 для 4х	66 для 3х, 75 для 4х	33 для 3 50 для 4 60 для 5		50	66	33-40
Стойкость при отказе диска	Нет	отличная		средняя		хорошая	отличная	хорошая	отличная
Восстановление данных		очень быстрое		быстрое		очень быстрое	быстрое	очень быстрое	среднее быстрое
Скорость случайного чтения			хорошая	очень хорошая			очень хорошая		
Скорость случайной записи	очень хорошая	хорошая	плохая	нормальная	плохая		хорошая		
Скорость линейного чтения				очень хорошая	хорошая		очень хорошая		
Скорость линейной записи				хорошая	средняя		хорошая		