

# вред и польза RAID-массивов

крик касперски ака мышьх, no-email

лет десять назад RAID-массивы были экзотикой, встречающейся только на серверах и крутых рабочих станциях, сейчас же RAID-контроллеры настолько подешевели, что массово встраиваются в материнские платы, соблазня юзеров объединить несколько дисков в один массив, но за это искушение приходится дорого платить

## введение

Смачная аббревиатура **RAID** расшифровывается как *Redundant Array of Independent | Inexpensive Disks* (Дисковый Массив Недорогих | Независимых Дисков) и появились эти штуки еще во времена "Железной Эры", когда землей правили динозавры, а на компьютерном рынке безраздельно господствовала фирма IBM.

Своим рождением они обязаны исследованием Петтерсона (David A. Patterson), Гибсона (Garth A. Gibson) и Катца (Randy H. Katz), предложивших для борьбы с отказами винчестеров (тогда они случались намного чаще, чем сейчас) использовать избыточный массив недорогих дисков (т. е. redundant array of inexpensive disks, или, сокращенно RAID) и запатентовавших эту идею в 1987 году (U.S. Patent 4,092,732 — <http://patft.uspto.gov/netacgi/nph-Parser?patentnumber=4,092,732>).

Идея получила развитие и очень скоро в RAID'ы стали объединять дорогие диски, поэтому слово "inexpensive" (дешевый) заменили на "independent" (независимый) и сейчас термин RAID принято расшифровывать именно так.

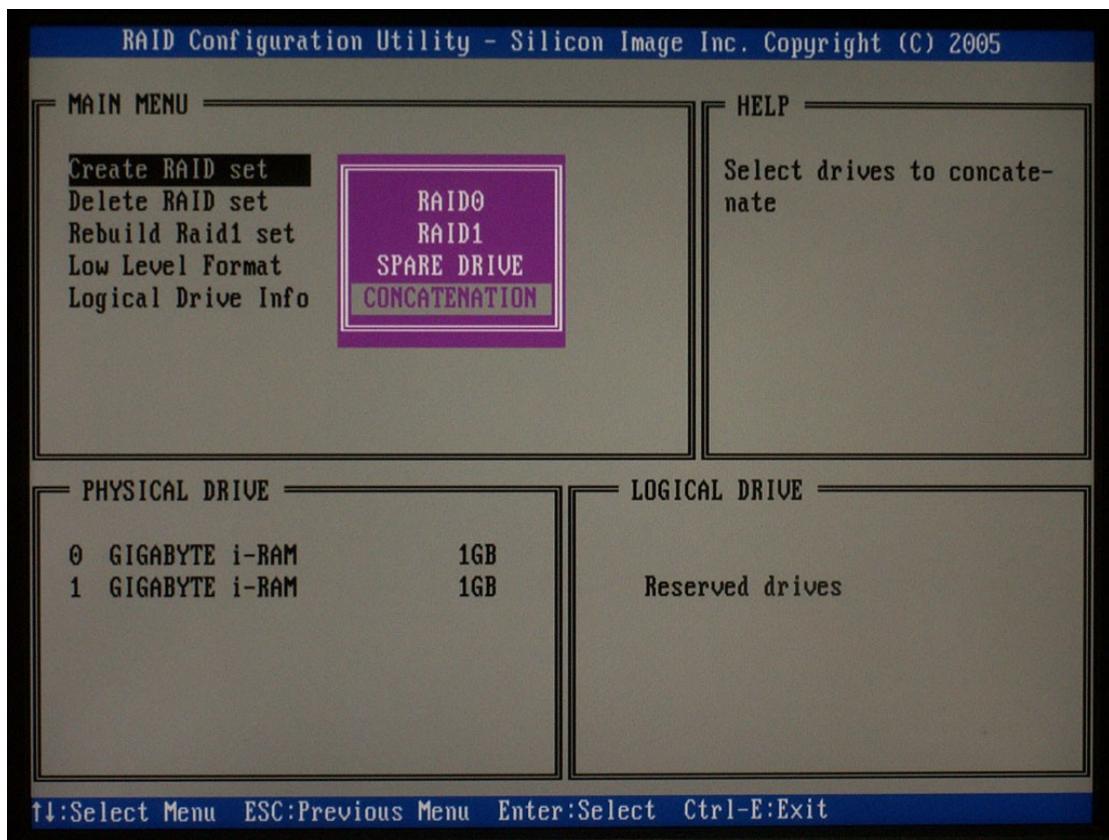


Рисунок 1 Adtron I35MB Diskpakt — RAID уровня 1 (Mirroring) из двух дисков с интегрированным контроллером, в форм-факторе 3.5" Single Slot

## **типы RAID'ов**

Что такое RAID? То, что массив — это и мышь'у понятно. Но вот как этим массивом можно распорядиться?! Если объединить несколько физических диска в один "виртуальный" и поблочно писать/читать со всех дисков сразу мы получим RAID уровня 0 (Striped Set), значительно увеличивающий скорость операций ввода/вывода (а, как известно, в большинстве случаев, ввод/вывод — самое узкое место в системе). В грубом приближении скорость обмена прямо пропорциональна количеству дисков, т.е. два диска увеличивают производительность практически вдвое.

Однако, здесь есть одно "но". Время поиска секторов (особенно при "далеких" позиционированиях) непостоянно и винчестер гоняет магнитную головку и в хвост, и в грибу, отыскивая нужный сектор методом вилки. При работе с одним жестким диском время поиска сектора равно  $X + n$ , где  $n$  — максимальное отклонение от среднего времени поиска, обусловленного конструктивными особенностями привода. При работе с массивом дисков время поиска увеличивается до  $X + n$ , поскольку контроллер вынужден дожидаться завершения операции обмена с самым медленным из дисков.



**Рисунок 2 выбор типа дискового массива в BIOS'e RAID-контроллера (не путать с системным BIOS'ом!)**

Как следствие — чем больше у нас дисков, тем меньший прирост производительности они дают. Переход с одного диска на два — чувствуется сразу, а добавление еще двух дисков уже практически не ощущается. Тем более, что при работе с мелкими файлами ожидаемого ускорения вообще не наступает, поскольку невозможно обеспечить перекрывающееся чтение/запись нескольких небольших сегментов информации.

К тому же, чем больше дисков, тем выше вероятность отказа каждого из них, а поскольку запись ведется блочным образом, т. е. блок 1 записывается на диск 1, блок 2 — на диск 2, блок 3 — на диск 3 и т. д., то при выходе одного диска из строя из данных образуется "решето", своеобразный реквием по хранящейся в RAID'e информации или, в переводе на русский язык — могила. RAID 0 — это крайне ненадежная штука и пользоваться ей можно только, например, для рабочей станции, предназначенной для обработчики цифрового видео, но ни в коем случае не для долговременного хранения данных!

RAID-контроллер				
диск 1	диск 2	диск 3	диск 4	диск 5
Сегмент 1	Сегмент 2	Сегмент 3	Сегмент 4	Сегмент 5
Сегмент 6	Сегмент 7	Сегмент 8	Сегмент 9	Сегмент 10

Таблица 1 схема хранения данных на RAID 0

А можно поступить иначе и писать одни и те же данные на все диски сразу, дублируя информацию, чтобы при выходе жесткого диска из строя, ее было возможно автоматически восстановить с другого. Такой массив называется "зеркальным" (Mirrored Set) и соответствует RAID уровню 1. Надежность резко возрастает, а время поиска секторов по прежнему равняется  $X + n$ , т.е. RAID 1 не только не ускоряет, но даже замедляет обмен данных по сравнению с обычным жестким диском.

RAID-контроллер				
диск 1 (данные)	диск 2 (копия диска 1)	диск 3 (данные)	диск 4 (копия диска 3)	диск 5 (свободный)
сегмент 1	сегмент 1	сегмент 2	сегмент 2	
сегмент 3	сегмент 3	сегмент 4	сегмент 4	

Таблица 2 схема хранения данных на RAID 1

RAID'ы уровней 2 и 3 обычно работают сразу с пятью дисками, храня на них не только полезные данные, но и байты четности, позволяющие восстановить вышедший из строя диск за счет остальных. Покажем как это происходит на примере битов (тем более, что восстановление данных есть битовая операция), пусть на четыре диска пишется: 1 0 1 1. Складывая биты друг с другом, получаем 3, а 3 по модулю 2 равно 1, вот этот самый 1 мы и пишем на пятый диск. Теперь, представим себе, что второй диск вышел из строя: 1 x 1 1 1. Мы вновь складываем биты, делим полученный результат по модулю 2 и получаем 0, что и требовалось доказать!

В принципе, дисков не обязательно должно быть пять, некоторые контроллеры работают и с тремя, но это неважно. Факт тот, что мы получаем и производительность, и надежность. Разница между RAID 2 и RAID 3 заключается в том, что RAID 2 использует для хранения битов четности несколько дисков, в то время как RAID 3 —только один, вследствие чего RAID 2 встречается крайне редко, да и RAID 3 используется нечасто.

RAID-контроллер				
диск 1 (данные)	диск 2 (данные)	диск 3 (данные)	диск 4 (данные)	диск 5 (информация четности)
байт 1	байт 2	байт 3	байт 4	байт четности
байт 5	байт 6	байт 7	байт 8	байт четности

Таблица 3 схема хранения данных на RAID 3

RAID уровня 4 — это фактически усовершенствованный RAID 3, устраняющий проблему производительности при работе с небольшими объемами информации, за счет использования только того диска, на котором эта информация хранится, что позволяет обрабатывать несколько запросов на чтение одновременно (класс! особенно при работе с кучей мелких файлов), однако, запросы на запись порождают блокировки, вызванные необходимостью обновления байтов четности и в настоящее время RAID 4 используется крайне редко.

RAID уровня 5 обходит проблему блокировок при записи путем распределения байтов четности по всем дискам, обеспечивая максимально возможную скорость обмена и высокую отказоустойчивость. Естественный недостаток — контроллеры этого типа довольно дорого стоят, к тому же как минимум нам потребуются три жестких диска, которые тоже денег стоят! Впрочем, цены как на диски, так и на контроллеры неуклонно снижаются и в обозримом будущем, RAID 5 рискует стать самым популярным типом из всех остальных.

RAID-контроллер				
диск 1	диск 2	диск 3	диск 4	диск 5
сегмент четности	сегмент 1	сегмент 2	сегмент 3	сегмент 4
сегмент 5	сегмент четности	сегмент 6	сегмент 7	сегмент 8
сегмент 9	сегмент 10	сегмент четности	сегмент 11	сегмент 12

**Таблица 4 схема хранения данных на RAID 5**

Большинство дешевых контроллеров, не поддерживающих RAID 5, тем не менее позволяют создавать гибридные дисковые массивы на основе RAID 0 + RAID 1, но в этом случае нам потребуется по меньшей мере четыре жестких диска, а избыточность будет составлять 50%, против 30% у RAID 5, работающего всего с тремя дисками.

Помимо аппаратных, существуют так же и программные RAID'ы, поддерживаемые практическими всеми современными операционными системами (такими как Linux/xBSD/NT/W2K/XP, а вот Windows 9x в этот список не входит). Будучи подключенными к различными IDE-каналам они обеспечивают схожую (или чуть-чуть более низкую) производительность, зато по гораздо более дешевой цене. С точки же зрения надежности, программный RAID ничем не отличается аппаратного, если не сказать, что выигрывает у него, поскольку чем больше железа установлено в компьютер, тем выше риск его поломки — это закон, против которого не попрешь!

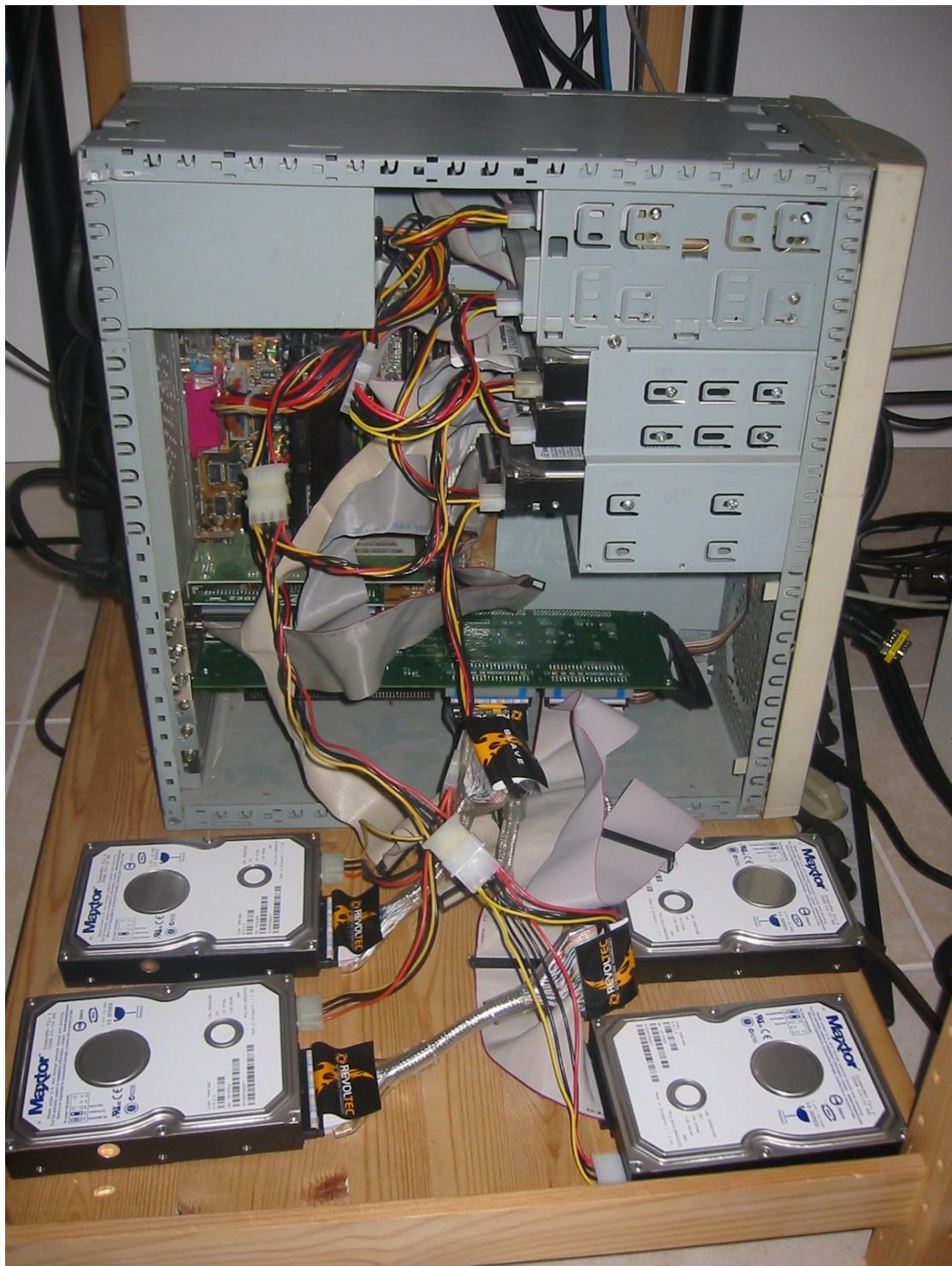


Рисунок 3 пяти дискам, объединенным в один RAID массив в бюджетном корпусе просто не хватило места!

## ПОЛЬЗА RAID МАССИВОВ

RAID'ы обеспечивают либо отказоустойчивость, либо производительность, либо и то, и другое сразу. Причем, на дисках, поддерживающих горячую замену – она же "hot-plug" (SCSI, SATA), выключать компьютер для замены вышедшего из строя винчестера необязательно. Это кстати говоря, актуально не только для серверов, но и рабочих станций, занимающихся обработкой цифрового видео или другими продолжительными расчетами, которые не могут

быть прерваны и которые занимают десятки часов машинного времени. Остановка системы означает, что все придется начинать заново и уже потраченные часы потрачены впустую.

Так же, объединение нескольких дисков в один, позволяет создавать разделы большого размера. Весьма актуально на RAID 0 при работе с морально устаревшими (а, значит, дешевыми) винчестерами на несколько гектар. Объединив их вместе, мы получим здоровенный раздел, доставшийся нам практически задаром.

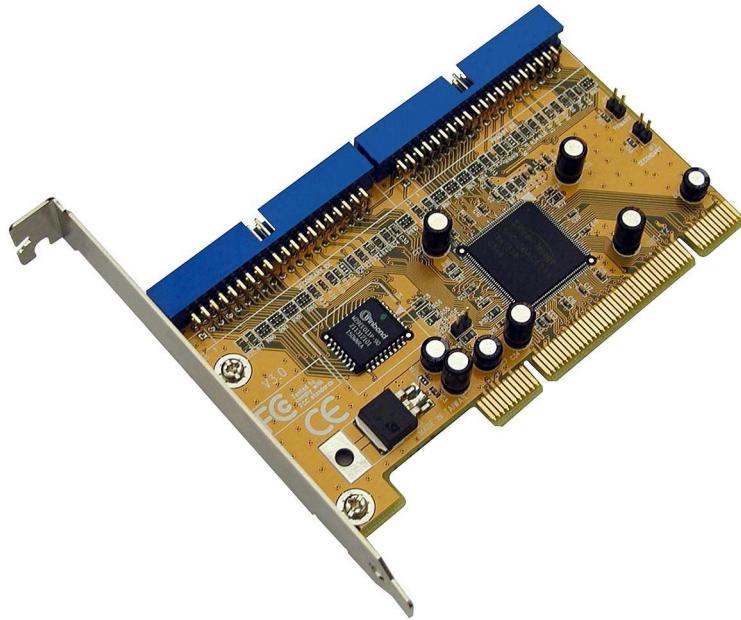
Вот, собственно, и все преимущества RAID'ом. Теперь проговорим об их недостатках.

## **...и вред**

RAID-массив страхует только от аппаратного выхода одного (реже — двух) жестких дисков из строя, но бессилен противостоять другим типам разрушения данных, как-то: вирусы, ошибки оператора, хакерские атаки, сбои операционной системы и ее окружения. А при выходе из строя блока питания или падения компьютера со стола (например), все диски массива обычно вылетают разом. Таким образом, вне зависимости от наличия/отсутствия RAID'a регулярное резервирование данных все равно остается обязательным! А если есть резервная копия, тогда какая польза от RAID'a?!

Восстанавливать же информацию с RAID'ов намного сложнее. Их поддерживают далеко не все "лечащие" утилиты, а фирмы, специализирующиеся на восстановлении — дерут за работу двойную - тройную цену, если вообще не отказываются от нее!

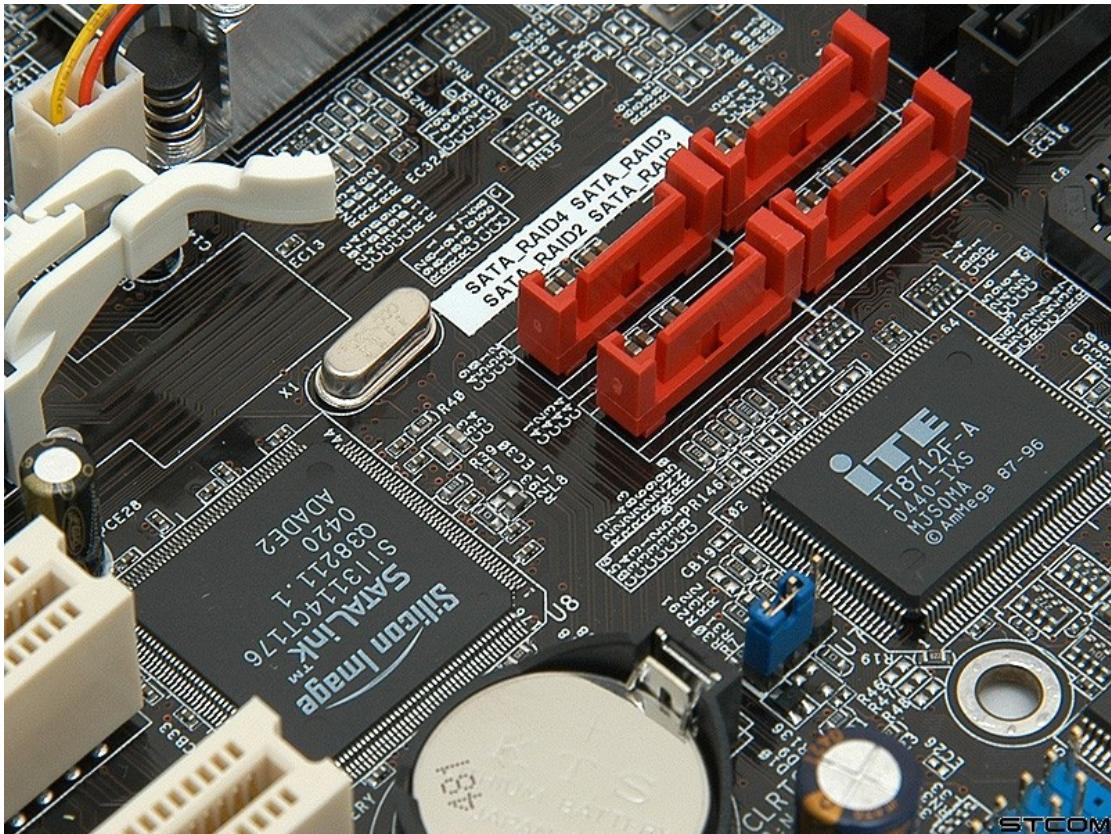
Причем, RAID-массив может работать только с тем контроллером, которым он был создан. Контроллеры различных производителей (и даже разные модели контроллеров одного и того же производителя) несовместимы друг с другом и если навернется контроллер, то... прости прощай весь дисковый массив! Хорошо, если RAID-контроллер внешний, тогда его, по крайней мере можно будет попытаться купить (хотя и это не всегда возможно, т.к. он уже давно можно быть снят с выпуска и предан забвению), а вот при выходе из строя RAID-контроллера на материнской плате (равно как и любого другого жизненно важного компонента материнской платы), наступает тырдец. Полный и окончательный. Потому что нам придется купить точно такую же материнскую плату, а их модельные ряды обновляются каждый сезон и "старички" исчезают из прайс-листов и складов со скоростью торнадо.



**Рисунок 4 Ultra ATA 133 PCI RAID-контроллер от никому не известной фирмы Maplin Electronics – штука хуже динамита**

Идея купить несколько внешних RAID-контроллеров "про запас", конечно хорошая (если не учитывать финансовую сторону дела), но глубоко неправильная, ибо современные микросхемы очень часто дохнут из-за сложных физико-химических процессов, разрушающих их изнутри. Чего стоит только один рост кристаллов в подложке, приводящих к утечкам проявляющихся себя от настырных сбоев до полной неработоспособности. Причем, поскольку нагрев и другие физико-механические воздействия разрушает кристаллы еще в зародыше

(точнее, тормозит их развитие), то контроллер, мирно покоящийся на полке имеет все шансы сдохнуть раньше, чем тот, что постоянно находится в работе.



**Рисунок 5 материнская плата с популярным интегрированным RAID-контроллером Silicon Image**

В этом смысле программные RAID-массивы более предпочтительны, поскольку работают с любым IDE/SCSI-контроллером, но и тут есть свои тонкости. Вплоть до Windows 2000, система хранила данные о дисковом массиве в... реестре! Следовательно, крах (или переустановка с нуля) Windows, означал и крах самого дискового массива со всеми содержащимися в нем данными. В Windows 2000 и более старших версиях, Microsoft высадила концепцию "динамических дисков", по сути дела, тех же самых программных RAID'ов, но! Теперь вся служебная информация хранится уже не в реестре, а на самом диске (дисках) и они могут быть подключены к любой другой W2K, XP.

А вот диски, подключенные к RAID-контроллеру (неважно — интегрированному или нет) в сумке уже не поносишь и просто так из одного компьютера в другой не перетыкнешь! Как же тогда обмениваться данными друг с другом? Как осуществлять апгрэд системы? Интегрированные контроллеры — привязывают нас к материнской плате, не интегрированные — к текущей операционные системы, поскольку вовсе не факт, что другая ось поддерживает данный RAID-контроллер, как не факт, что его производитель поддерживает все оси, которые нам нужны. Ну Windows еще туда-сюда, хотя на "устаревшие" (по мнению производителей) контроллеры новые драйвера как правило не выпускаются, со всеми вытекающими отсюда...

Наконец, у дисков, подключенных к RAID-контроллерам, очень сложно прочитать показания S.M.A.R.T. (системы самотестирования и мониторинга), а иногда и вообще невозможно. А читать их крайне полезно, поскольку по ним можно с некоторой вероятностью предсказать сколько еще винту жить осталось и узнать его температуру. Когда диск один — за его температуру можно, в общем-то, и не волноваться, но вот массив из 4x тесно расположенных дисков (а по другому их располагать даже в BigTower'e никак не получается) способен нагреваться до весьма высоких температур, требующих немедленной установки дополнительных систем охлаждения, иначе... Впрочем, не будем о грустном. Ведь позади новый год, а впереди весна, май. А май это трава. Так... о чём это мы с вами тут говорили?

## **советы по выбору RAID'a**

Если же вы все-таки решили объединить диски в RAID и никакими силами природы вас не отговорить, то по крайней мере остановите свой выбор на контроллере именитого производителя, такого, как, например, Adaptec, который назавтра не исчезнет вместе со своими контроллерами.

Что же касается интегрированных контроллеров, тут тоже следует выбирать наиболее популярные чипы, используемые многими производителями материнских плат (хотя и не факт, что "чужая" материнская плата даже с тем же самым чипом "подхватит" уже собранный RAID). Если на материнской плате имеется несколько контроллеров (например, от Intel, Integrated Technology Express Inc и Silicon Image), следует выбирать тот, что распознается операционной системой без установки внешних драйверов (в данном случае это Silicon Image), в противном случае при ремонте упавшей системы могут возникнуть серьезные проблемы самая распространенная из которых — грузимся с Live CD, а ось не видит RAID'a. В упор. Кстати говоря, Silicon Image штатное поддерживается не только NT, но и остальными операционными системами (типа Linux или xBSD), а вот Intel — увы. Integrated Technology Express Inc (известный под аббревиатурой ITE не поддерживается NT, но понимается Linux'ом и xBSD, так что в случае аварии KNOPPIX LiveCD нам поможет!)



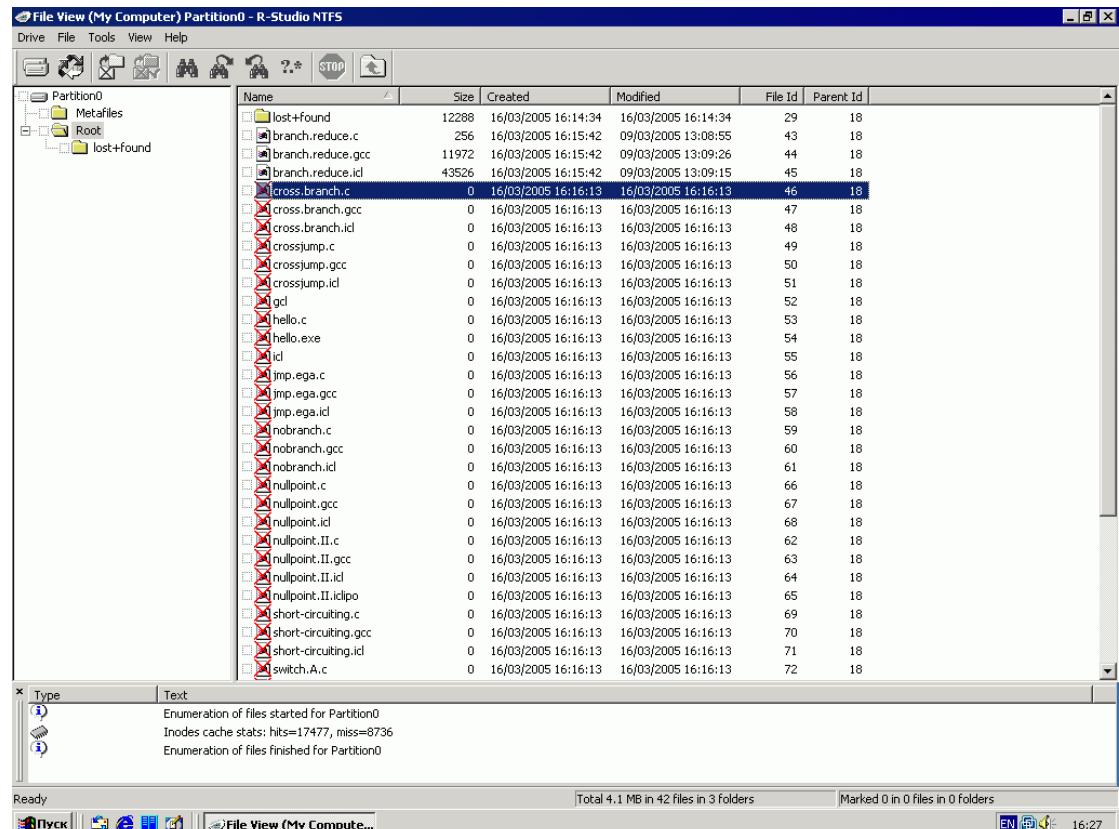
**Рисунок 6 чип от Silicon Image – сердце RAID-контроллера**

Существует нелепая легенда, что при установке в RAID требуются винчестеры одной модели или, по крайней мере, одного производителя. Что за чушь! Можно брать любых диски, не только от разных производителей, но даже разного размера. В случае RAID'a уровня 0 мы получим диск суммарной емкости, в случае RAID 1 – наименьшей из всех имеющихся. Вся хитрость в том, что, как неоднократно показывала практика, диски одной модели, имеющие

дефекты проектирования или "косяки" в техпроцессе, дохнут примерно в одно и тоже время с разбросом всего в несколько месяцев. Подобный массовый падеж на RAID'ах приводит к тому, что члены массива выходят из строя один за другим прежде, чем владелец машины успевает заменить их.

Выбор дисков от разных производителей \_реально\_ увеличивает надежность хранения данных на RAID 1 и RAID 5. Что же касается RAID 0, то даже при отказе одного из дисков, он теряет все хранимые на нем данные и тут, действительно, лучше выбирать идентичные модели, обладающие сходными скоростными характеристиками.

И последнее. Если RAID все-таки упал, не паникуйте и не рвите на себе зубами вены. Возьмите R-Studio от R-Tools Technology (<http://www.r-tt.com/>) – это лучшее средство для автоматического восстановления из всех имеющихся.



**Рисунок 7** внешний вид утилиты R-Studio от R-Tools Technology, умеющей восстанавливать разрушенные RAID-массивы в автоматическом режиме

## заключение

Так что же, все-таки, по большому счету представляют собой RAID'ы? Добро или зло? Вопрос сложный, можно даже сказать риторический. И как к любому риторическому вопросу к нему надо подходить к философской точки зрения. То есть выбирать, руководствуясь собственной интуицией и чутьем.