

УДК 004.65

К ВОПРОСУ ВЫБОРА НАИЛУЧШЕГО УРОВНЯ RAID ДЛЯ ХРАНИЛИЩ ДАННЫХ ИНФОРМАЦИОННОЙ СИСТЕМЫ, ОБЕСПЕЧИВАЮЩЕЙ БЫСТРУЮ ОБРАБОТКУ БОЛЬШИХ ДАННЫХ

Атрощенко В.А., Тымчук А.И.

ФГБОУ ВО «Кубанский государственный технологический университет», Краснодар,
e-mail: atros.v@yandex.ru

Данная статья посвящена вопросу выбора наилучшего уровня RAID для хранилищ данных информационной системы, обеспечивающей быструю обработку больших данных. Поскольку обработка больших данных – ресурсоёмкий процесс, а ценность информации, полученной при анализе больших данных, может быть весьма высокой, то для хранилищ данных системы необходим RAID, обеспечивающий их высокую производительность и высокую вероятность безотказной работы. В статье рассматриваются уровни RAID, которые можно построить с помощью аппаратных контроллеров и архитектура которых не является проприетарной. Для определения наилучшего из исследуемых уровней RAID для каждого из них строится функция оптимальности, зависящая от трёх критериев – вероятность безотказной работы, производительность, эффективность использования дискового пространства, с учётом ограничений, определяющих их важность. Также в связи с ростом вероятности отказа массива при увеличении числа используемых дисков вводится ограничение на их максимальное число. После рассмотрения графиков функций оптимальности на отрезке, ограниченном максимальным числом используемых дисков, делается вывод, что RAID 10 на рассматриваемом отрезке, является наилучшим, для использования в хранилищах данных информационной системы, обеспечивающей быструю обработку больших данных.

Ключевые слова: большие данные, хранилища данных, RAID

ON THE QUESTION OF CHOOSING THE BEST RAID LEVEL FOR DATA STORAGE OF THE INFORMATION SYSTEM THAT PROVIDES FAST PROCESSING OF BIG DATA

Atroschenko V.A., Tymchuk A.I.

Federal Budget Educational Institution of Higher Education Kuban State Technological University,
Krasnodar, e-mail: atros.v@yandex.ru

This article deals with the issue of choosing the best RAID level for data warehouses of the information system that provides fast processing of large data. Since the processing of big data is a resource-intensive process, and the value of the information obtained in the analysis of big data can be very high, the system data warehouses need the RAID, providing its high performance and a high probability of failure-free operation. The article deals with the RAID levels, which can be constructed using hardware controllers, and the architecture of which is not proprietary. To determine the best from studied RAID levels for each of them is constructed a function of optimality that depends on three criteria – the probability of failure-free operation, performance, efficient use of disk space, taking into account the constraints that define their importance. Also because of increasing, the probability of failure of the array while increasing numbers of discs used by it, administered at their maximum number limit. After considering graphics of optimality functions on the interval limited by the maximum number of drives used, it is concluded, that a RAID 10 on the given interval is the best for use in the data storages of information system that provides fast processing of big data.

Keywords: big data, data storages, RAID

Одна из основных функций систем хранения данных – это обеспечение сохранности информации и предотвращение её утраты. Информация может быть утеряна по разным причинам, но чаще всего, если не принимать никаких мер для предотвращения, это происходит при выходе из строя устройств хранения данных (жёстких дисков, серверов и т.д.). По данным статистики вероятность поломки жёсткого диска в первый год равна приблизительно 1,3% для дисков desktop класса, 1% для near line класса и 0,5% для enterprise класса [5]. В крупных системах хранения данных число используемых жёстких дисков может исчисляться сотнями и тысячами, что увели-

чивает вероятность выхода из строя хотя бы одного из дисков пропорционально их количеству. Так, например, для группы из 100 дисков near line класса ежегодная вероятность выхода из строя в первый год эксплуатации хотя бы одного будет равна 100%, что приводит к необходимости использования решений для предотвращения утраты информации стоимость которой может намного превышать стоимость её хранилищ.

Основным решением для данной проблемы является использование RAID технологии виртуализации данных, которая объединяет несколько дисков в логический элемент для избыточности и повышения производительности [9]. Существуют раз-

личные уровни RAID – архитектуры их построения: RAID 0 (чередование), RAID 1 (зеркалирование), RAID 2 (использование кодов Хемминга), RAID 3 (использование диска чётности, информация разбита на байты), RAID 4 (диск чётности, информация разбита на блоки), RAID 5 (информация разбита по всем дискам массива), RAID 6 (дублирование чётности), а также их различные комбинации, например RAID 10 – комбинация RAID 1 и RAID 0 [1]. Они в свою очередь делятся на три различных подхода к реализации по типу RAID контроллера:

1) программный RAID – реализуемый при помощи программного обеспечения и использующий для управления массивом ресурсы ЦП [10];

2) аппаратный RAID – плата расширение или внешнее устройство, применяющие для управления массивом свои собственные процессоры и не использующие ресурсы ЦП;

3) полуаппаратный (интегрированный в материнскую плату) RAID – отдельный чип, встроенный в материнскую плату, выполняющий функции управления массивом, но при этом также использующий ресурсы ЦП.

Выбор уровня и подхода к реализации различной архитектуры RAID зависит от конкретной системы хранения данных, для которой он создаётся.

Для хранения данных информационной системы, обеспечивающей быструю обработку больших данных, необходим RAID с максимально возможной вероятностью безотказной работы и производительностью, в пределах заданной стоимости системы.

В соответствии с поставленными условиями для построения RAID лучшими являются аппаратные контроллеры, поскольку, во-первых, при одинаково высокой отказоустойчивости их производительность выше, чем у программных, за счёт того, что они не используют ресурсы серверного процессора, а во-вторых, благодаря использованию собственного процессора могут использовать кэш-память, что положительно сказывается на производительности массива, в особенности при проведении операций записи данных [3].

Для определения наилучшего из уровней RAID в соответствии с этими же условиями необходимо сначала рассмотреть их основные характеристики, представленные в таблице [2].

Из таблицы видно, что производительность массива рассчитывается как среднее арифметическое между его IOPS чтения и IOPS записи. Экономичность напрямую зависит от эффективности использования дискового пространства. Вероятность безотказной работы рассчитывается в соответствии с формулой Бернулли на основе показателя максимально допустимого числа вышедших из строя дисков, при котором массив продолжает работу.

Для одиночного диска вероятность безотказной работы N равна $1 - S$, где S – вероятность отказа среднестатистического одиночного жёсткого диска класса в результате износа в соответствии с нормальным законом распределения в течение первого года эксплуатации, равная 0,01 при работе 24 часа в сутки 365 дней в году [4].

Основные характеристики уровней RAID

Основные характеристики исследуемых уровней RAID					
RAID	К	Е	Х	Р	W
0	От 2	К	0	I*K	I*K
1	Кратно 2	К/2	От 1 до К/2	I*2	I/2
5	От 3	К-1	1	I*K	I*K/4
6	От 4	К-2	2	I*K	I*K/6
10	От 4, чётное	К/2	От 1 до К/2	I*K	I*K/2
50	От 6, чётное	К-2	От 1 до 2	I*K	I*K/4
51	От 6, чётное	(К-1)/2	От 2 до К/2+1	I*K	I*K/8
60	От 8, чётное	К-4	От 2 до 4	I*K	I*K/6
61	От 8, чётное	(К-4)/2	От 4 до К/2+2	I*K	I*K/12

Примечание. К – число дисков, требуемое для создания массива. Е – эффективно используемое число дисков (число дисков из общего количества дисков массива, не использующееся для создания избыточности). Х – максимально допустимое число вышедших из строя дисков, при котором массив продолжает работу и возможно восстановление хранимой в нём информации. Р – IOPS (Input/Output Operations Per Second – количество операций ввода/вывода в секунду) [7]. чтения массива [11]. W – IOPS записи массива [11]. I – IOPS одиночного жёсткого диска, аналогичного жёстким дискам, используемым в массиве.

Массив RAID 0 будет продолжать работать пока работают все его диски, следовательно

$$N(\text{RAID } 0) = (1 - S)^K;$$

Массив RAID 1 работоспособен, пока в каждой паре работают один или оба диска, поэтому его вероятность безотказной работы равна

$$N(\text{RAID } 1) = \\ = \left(2(1 - S)S + (1 - S)^2\right)^{\frac{K}{2}} = (1 - S^2)^{\frac{K}{2}}.$$

Массив RAID 5 работоспособен, пока работают все его диски или все кроме одного любого из них:

$$N(\text{RAID } 5) = (1 - S)^K + K(1 - S)^{K-1} S.$$

$$N(\text{RAID } 10) = \left(2(1 - S)S + (1 - S)^2\right)^{\frac{K}{2}} = (1 - S^2)^{\frac{K}{2}};$$

$$N(\text{RAID } 50) = \left((1 - S)^{\frac{K}{2}} + \frac{K}{2}(1 - S)^{\frac{K}{2}-1} S\right)^2;$$

$$N(\text{RAID } 60) = \left((1 - S)^{\frac{K}{2}} + \frac{K}{2}(1 - S)^{\frac{K}{2}-1} S + 0,25K(0,5K - 1)(1 - S)^{\frac{K}{2}-2} S^2\right)^2.$$

RAID 51 это два зеркальных RAID 5, поэтому вероятность безотказной работы RAID 51 получается подстановкой в $N(\text{RAID } 1)$ при $K = 2$ вместо S вероятности отказа RAID 5 равную $1 - N(\text{RAID } 5)$.

$$N(\text{RAID } 51) = 1 - \left(1 - (1 - S)^{\frac{K}{2}} + \frac{K}{2}(1 - S)^{\frac{K}{2}-1} S\right)^2;$$

Аналогичным образом рассчитывается вероятность безотказной работы для массива RAID 61, являющегося парой зеркальных RAID 6.

$$N(\text{RAID } 61) = 1 - \left(1 - (1 - S)^{\frac{K}{2}} + 0,5K(1 - S)^{\frac{K}{2}-1} S + 0,25K(0,5K - 1)(1 - S)^{\frac{K}{2}-2} S^2\right)^2.$$

Для определения наилучшего из исследуемых уровней RAID используем функцию оптимальности, зависящую от этих критериев, с учётом ограничений, определяющих важность критериев – вероятность безотказной работы, производительность, эффективность использования дискового пространства при использовании пятибалльной шкалы будет равна 5, 5, 1. Коэффициенты для них соответственно будут равны: $\frac{5}{11}$, $\frac{5}{11}$ и $\frac{1}{11}$. Поскольку критерии имеют различные единицы и масштабы измерения, их необходимо нормировать. При этом функция оптимальности будет иметь следующий вид:

Аналогичным образом рассчитывается вероятность безотказной работы для массива RAID 6, который сохраняет работоспособность при потере не более двух любых своих жёстких дисков:

$$N(\text{RAID } 6) = \\ = (1 - S)^K + K(1 - S)^{K-1} S + \\ + 0,5K(K - 1)(1 - S)^{K-2} S^2.$$

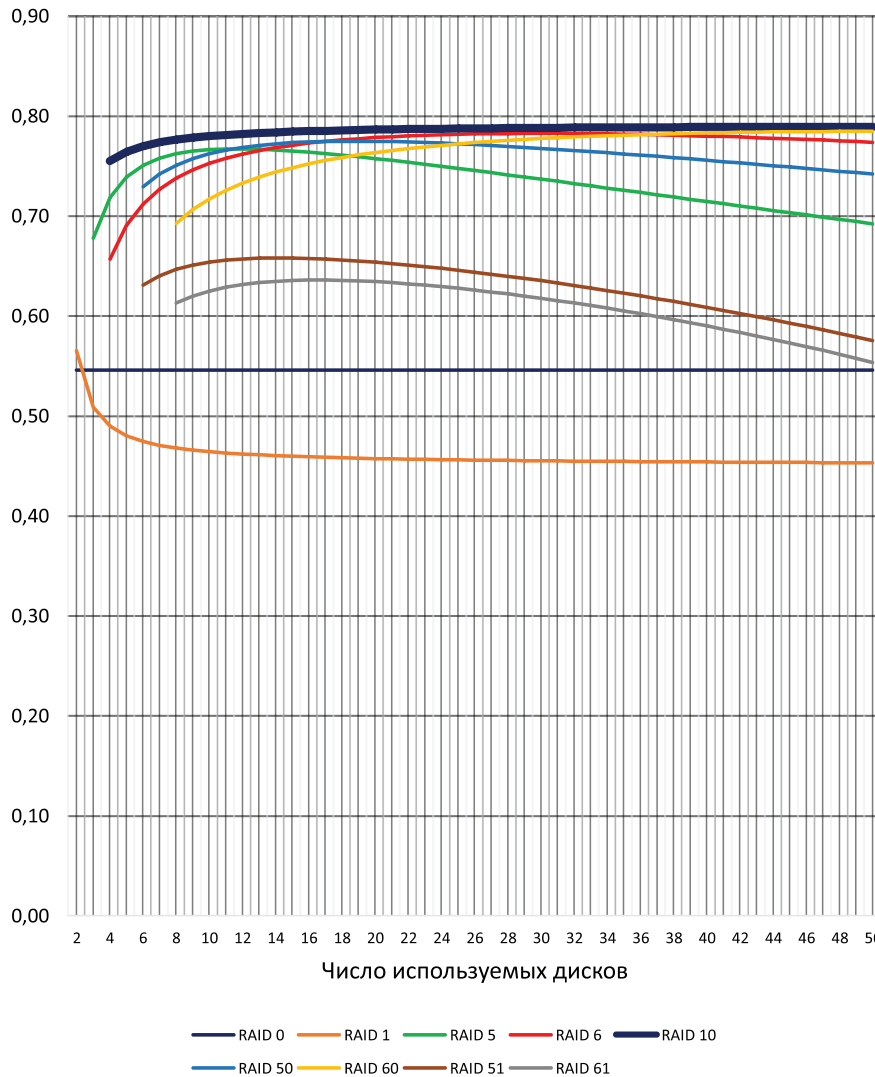
Вероятность безотказной работы RAID 10, RAID 50 и RAID 60 получается подстановкой в $N(\text{RAID } 0)$ вместо вероятности выхода из строя одиночного диска вероятности выхода из строя массива RAID 1 для $\frac{K}{2}$ зеркал, RAID 5 и RAID 6 из $\frac{K}{2}$ дисков соответственно.

$$f(K) = \frac{5}{11} * \frac{P_{RAID}(K) - P_{min}(K)}{P_{max}(K) - P_{min}(K)} + \\ + \frac{1}{11} * \frac{E_{RAID}(K) - E_{min}(K)}{E_{max}(K) - E_{min}(K)} + \\ + \frac{5}{11} * \frac{N_{RAID}(K) - N_{min}(K)}{N_{max}(K) - N_{min}(K)},$$

где $P_{RAID}(K)$ – производительность RAID в IOPS равная среднему арифметическому IOPS чтения и IOPS записи уровня RAID.

$P_{max}(K)$ – максимальная производительность, которой можно добиться, используя K дисков. $P_{max}(K) = P_{RAID 0}(K)$.

Графики значений функции $f(K)$ для исследуемых уровней RAID



Графики функций оптимальности для исследуемых уровней RAID

$P_{\min}(K)$ – минимальная производительность, которой можно добиться, используя K дисков. $P_{\max}(K) = P_{JBOD}(K)$. JBOD (Just a bunch of disks) – дисковый массив, в котором единое логическое пространство распределено по жёстким дискам последовательно [8].

$E_{RAID}(K)$ – эффективно используемое число дисков в RAID массиве, состоящем из K дисков.

$E_{\max}(K)$ – максимально возможное число эффективно используемых дисков из K доступных. $E_{\max}(K) = E_{RAID0}(K)$.

$E_{\min}(K)$ – минимально возможное число эффективно используемых дисков из K доступных. $E_{\min}(K) = E_{RAID1}(K)$.

$N_{RAID}(K)$ – вероятность безотказной работы RAID массива, состоящего из K жёстких дисков.

$N_{\max}(K)$ – максимально возможная вероятность безотказной работы группы из K дисков. $N_{\max}(K) = 1$.

$N_{\min}(K)$ – минимально возможная вероятность безотказной работы группы из K дисков. $N_{\min}(K) = N_{JBOD}(K)$.

Поскольку с увеличением числа дисков в массиве растёт вероятность его отказа по различным причинам и потери хранящейся в нём информации, то появляется необходимость в введении ещё одного ограничения на число дисков, входящих в массив. Поэтому будем рассматривать RAID мас-

сивы, включающие в себя не более 50 дисков. Графики функций оптимальности $f(K)$ для исследуемых уровней RAID при этом, как показано на рисунке, имеют следующий вид.

Как видно из графиков функций оптимальности для исследуемых уровней RAID, представленных на рисунке, на исследуемом промежутке числа используемых дисков K , значения функции оптимальности для RAID 10 являются максимальными по сравнению со значениями функций оптимальности остальных исследуемых уровней RAID для любого значения K , за счёт наилучшей, после RAID 0, производительности среди всех исследуемых уровней RAID и высокой вероятности безотказной работы, уступающей только аналогичному показателю RAID 6. Остальные исследуемые уровни RAID, построенные на базе первого уровня и сам RAID 1 серьёзно проигрывают десятому по производительности, а построенные на базе пятого с увеличением числа используемых жёстких дисков серьёзно проигрывают ему в вероятности безотказной работы.

Однако также можно заметить, что при числе дисков от 24 до 32 значения функции оптимальности для RAID 6, а при числе дисков больше 30 значения функции оптимальности для RAID 60 практически равны значениям аналогичной функции для RAID 10, что при использовании данного числа дисков делает их достойной альтернативой RAID 10 при необходимости увеличить эффективность использования дискового пространства массива и сократить тем самым расходы на его создание.

Помимо того, что RAID 10 является наилучшим с точки зрения принятой функции оптимальности, он имеет также ряд других преимуществ перед большинством исследуемых уровней RAID. Во-первых, он имеет преимущество в скорости восстановления данных в случае выхода из строя одного из жёстких дисков массива и снижении при этом его производительности [6], поскольку в RAID 10, как и в RAID 1, для восстановления данных будет задействован только один диск зеркальный потерянного, в то время как в остальных исследуемых уровнях RAID, кроме RAID 0, где восстановление вообще невозможно, для этого будут использованы все диски массива, что на время необходимое для восстановления массива создаёт дополнительную нагрузку на систему и существенно снижает скорость её работы.

Вторым преимуществом RAID 10 является то, что вероятность выхода его из строя в процессе восстановления, из-за

UER (unrecoverable error rate) – вероятности появления невосстановимой ошибки чтения на жёстком диске, по различным причинам: дефект поверхности, сбой в работе головки, контроллера и т.д. [5], – равна аналогичной у RAID 1 и меньше чем у любого другого уровня RAID. Этим преимуществом он таже, как и первым, обязан тому обстоятельству, что при восстановлении потерянного жёсткого диска используется лишь зеркальный ему жёсткий диск, а не все остальные диски массива, поскольку вероятность невосстановимой потери данных массивом в процессе восстановления по причине UER прямо пропорциональна числу жёстких дисков, которые задействованы в этом процессе.

Подводя итог, можно сделать вывод, что наилучшим RAID массивом для использования в хранилищах данных информационной системы, обеспечивающей быструю обработку больших данных, является массив с архитектурой RAID 10, построенный на базе аппаратного RAID контроллера, поскольку он обеспечивает наилучшее соотношение между высокой производительностью и высокой вероятностью безотказной работы.

Список литературы

1. Введение в raid: основные термины и подходы [Электронный ресурс]. – URL: <http://www.8host.com/blog/vvedenie-v-raid-osnovnye-termyny-i-podhody> (дата обращения: 15.12.2016).
2. Виды RAID и их характеристики [Электронный ресурс]. – URL: <http://pyatilistnik.org/vidyi-raid-i-ih-harakteristiki> (дата обращения: 21.12.2016).
3. Всё, что вы хотели узнать о RAID-контроллерах, но лень было искать [Электронный ресурс]. – URL: <http://habrahabr.ru/company/pc-administrator/blog/304798> (дата обращения: 25.12.2016).
4. Как установить надёжность моего RAID-массив [Электронный ресурс]. – URL: http://protonpc.ru/info/kak-ustanovit-nadezhnost-raid_60 (дата обращения: 26.12.2016).
5. Надёжность жёстких дисков [Электронный ресурс] – URL: <http://true-system.blogspot.ru/2013/04/mtbf-afr-uer-raid.html> (дата обращения: 07.12.2016).
6. Сравнение типов RAID, используемых в серверах и СХД [Электронный ресурс]. – URL: <http://altastor.ru/tech/raid/> (дата обращения: 29.03.2017).
7. Что такое IOPS и что он показывает [Электронный ресурс]. – URL: <http://pc-hard.ru/hardarticles/87-what-is-iops.html> (дата обращения: 29.03.2017).
8. Что такое JBOD [Электронный ресурс]. – URL: <http://pyatilistnik.org/chto-takoe-jbod> (дата обращения: 29.03.2017).
9. RAID – Википедия [Электронный ресурс]. – URL: <http://ru.wikipedia.org/wiki/RAID> (дата обращения: 05.12.2016).
10. RAID [Электронный ресурс]. – URL: http://wiki.vspru.ru/workroom/pi21-2008/что_такое_raid (дата обращения: 15.12.2016).
11. RAID Penalty [Электронный ресурс]. – URL: <http://theithollow.com/2012/03/21/understanding-raid-penalty> (дата обращения: 19.12.2016).