# RAID Configuration: Pragmatic Selection Strategies

Marcel Schubert

30.03.2024

## Introduction

Choosing the right configuration for the logical disks of your systems is no trivial task. But it is still one, many technicians are faced with.

The goal of this report is to present pragmatic aproaches on how to choose the right configuration for the specific task at hand. It does not include explanations on the underlying technical details. It also does only apply a narrow view, and does not consider other factors like access protocols or optimized implementations. But still the information can be used to calculate and compare theoretical values to achieve more informed decisions or to do further research.

The methology used is simply the analysis of available information online. Since this is not very scientific, most information was not checked beyond plain plausibility. Also note that i did not make any quantifiable empirical measurements.

The report is divided into a few main parts:

In the first part general methods to compare economical differences are presented.

In the second some methods are shown which can help calculate the theoretical performance of arrays. It also includes other factors which can be considered.

In the last part a general checklist is presented which can be used to quickly tackle the problem of finding an more or less optimal solution.

# Contents

# 1 Reliability

Since reallife situation depend on many variables simple numeric methods are not really conlusive. But consider the average failure time of one drive and then replicate it among a storage group, the probability of a drive failing will obviously increase drastically with higher numbers of drives. Analitically this change is hard to define, since it depents on many factors. For example: it is very common, that drives coming from the same batch of production might fail togheter in short succession, factors like these are hard to determine. [1]

Still, if one wants to get a simple overview further i will include formulas for the failure calculation for RAID-5 and RAID-6.

$$\frac{MTTF^2(disk)}{N * (G-1) * MTTR(disk)} \tag{1}$$

Equation 1 gives a way to calculate the mean time to failure of RAID-5 arrays. Where $MTTF(disk)$ is the mean time of failure of a single disk $N$ is the number of disks and $G$ is the size of the arrays. The equation assumes no correlated failures, that means that this simple model assume all disks are independent. The same calculation looks slightly different for RAID-6 as seen in eqation 2. [1]

$$\frac{MTTF^3(disk)}{N * (G-1) * (G-2) * MTTR^2(disk)} \tag{2}$$

A overview of the amount of hardware that can fail without disrupting service can be seen in table 1. Notice, that in RAID-1 it also very much depends on luck where the failure happens, since if the stripes are favourably replicated multiple drives might be able to fail. [2]

| **Level** | Failures |
|---|---|
| 0 | 0 |
| 1 | 1 or $\frac{n}{2}$ |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 2 |

Table 1: Amount of drives able to fail without array service degradation [2]

## 2 Capacity

A comparative capacity calculation can relatively easily be done using table 2.

| **Level** | Space Efficiency |
|---|---|
| 0 | n |
| 1 | $\frac{n}{2}$ |
| 2 | $(1 - \frac{1}{n} * log_2(n+1)) * full_caption$ |
| 3 | n-1 |
| 4 | n-1 |
| 5 | n-1 |
| 6 | n-2 |

Table 2: Coefficient of the space multiplication (smallest disk). [2] The calculation of RAID-2 was derived from Chen et al., Alagappan and wikipedia. [1][2]

# 3 Cost

A economic comparison can be created using the matrix shown in table 3. It is feasable to compare the cost point relative to RAID level-0, since it is the configuration with the lowest cost/efficiency rating. To compare the different options use $N =$ Number of disks and with $\max(x, y)$ is the known max function with $x, y \in R$. Small here refers to I/O requests of one striping unit, large refers to I/O requests of one full stripe (one stripe unit from each disk in an error- correction group). A general overview can be seen in figure 3 which plots this againts arrays of different sizes. [1]

| **Level** | Small Read | Small Write | Large Read | Large Write | Storage Efficiency |
|---|---|---|---|---|---|
| 0 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | $\frac{1}{2}$ | 1 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| 3 | $\frac{1}{N}$ | $\frac{1}{N}$ | $\frac{N-1}{N}$ | $\frac{N-1}{N}$ | $\frac{N-1}{N}$ |
| 5 | 1 | $\max\left(\frac{1}{N}, \frac{1}{4}\right)$ | 1 | $\frac{N-1}{N}$ | $\frac{N-1}{N}$ |
| 6 | 1 | $\max\left(\frac{1}{N}, \frac{1}{6}\right)$ | 1 | $\frac{N-2}{N}$ | $\frac{N-2}{N}$ |

Table 3: Cost Throughput Comparison relative to RAID-0 [1]

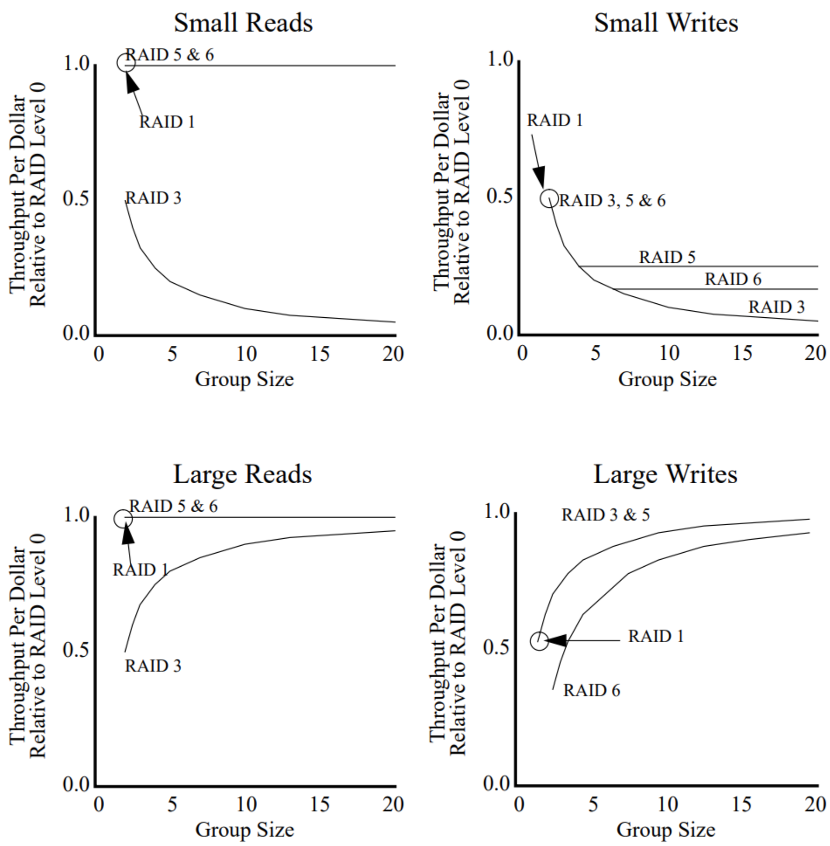


Figure 1: Cost Throughput Comparison

# 4 Performance

You can approximate theoretical upper bound of the performance by using table 4. The calculation only take conceptual first order factors into account. But despite this, it may still be benefitial to compare these fundamental boundaries to get a general overview. The table offers the coefficients which can be multiplied with the given drive metrics (random read, random write, seq read, seq write). Notice, that this does not use specific block sizes, it can be applied on average or median values. [2]

## 4.1 Nested RAID Configuraitions

For nested RAID configuration, calculate the values of a single group first and then iterate the calculation. Evenly distrubute the group size across the total amount of disks. For example: calculate the table then take the values of the RAID-5 configuration an plug it into the table again. Consider that this does not include many optimizations typically implemented on plattforms. Again keep in mind, that this is a rough approximation.

| **Level** | sread | swrite | rread | rwrite | rlatency | wlatency |
|---|---|---|---|---|---|---|
| 0 | n | n | n | n | 1 | 1 |
| 1 | $\frac{n}{2}$ | $\frac{n}{2}$ | n | $\frac{n}{2}$ | 1 | 1 |
| 4 | n-1 | n-1 | n-1 | $\frac{1}{2}$ | 1 | 2 |
| 5 | n-1 | n-1 | n | $\frac{n}{4}$ | 1 | 2 |
| 6 | n-2 [1] | n-2 [1] | n [1] | $\frac{n}{6}$ [1] | 1 [1] | 2 [1] |

Table 4: Performance Calculation [2], The original talbe does not include RAID-6 so it was derived from Chen et al [1]

Notice, that with increasing number of disks the upper bound of the throughput can be enhanced, but generaly latency is not improved.

## 4.2 Other Factors

Always read the specific hardware as well as software manuals and documents since theoratical speeds very much depend on the specific implementations. Factors like controller, algorithm and cache optimization play a important role. In the most general case, we just assume that they scale with a constant value.

# 5 Guide

For a fast comparison use the Microsoft Excel workbook "raid-workbook.xlsx" supplied with this report. Plug in the dependable variables like drive size and io speeds and further analyze the results.

## 5.1 HPE Smart Array

HP offers different storage solutions, one of these consists of the Smart Array Controllers, which offer different functionalities separated by classes. In this chapter i want to give the reader an overview and present, what HP writes about the topics reliability, performance and efficiency.

There are three main classes: S,E and P. S-Class provides a typical software RAID for MS Windows environments and is an entry level product. Next in line is the E-Class physical controllers which offer enterprise level functionalities but not all high performace optimiziations like caching. Last in line are P-Class physical controllers which offer typical performance oriented functionalities and optimizations like chaching and different types of interfaces. For the exact feature matrix consult the documentation found online. [3]

Performance whise HP suggests to consider, that the performance decreases as fault tolerance improves due to extra I/O as well as that read performance is generally the same for all RAID levels except for smaller RAID 5 or 6 arrays. [3]

In their documentations they almost show the same theoretical numbers to create a performance comparison. They simplify their model on the basis of needed write io operations.

- RAID-0, 1
- RAID-1/10, 2
- RAID-1/10 Triple, 3
- RAID-5, 4
- RAID-6, 6

This has evidently a big enough similiarity to the model used in chapter "Performance".

For the capacity aspects, they suggest to consider, that usable capacity decreases as fault tolerance improves due to an increase in parity data. Further they say, that the usable capacity for RAID 10 and RAID 10 Triple remains flat with larger arrays and that the usable capacity for RAID 5, 50, 6, and 60 increases with larger arrays. [3]

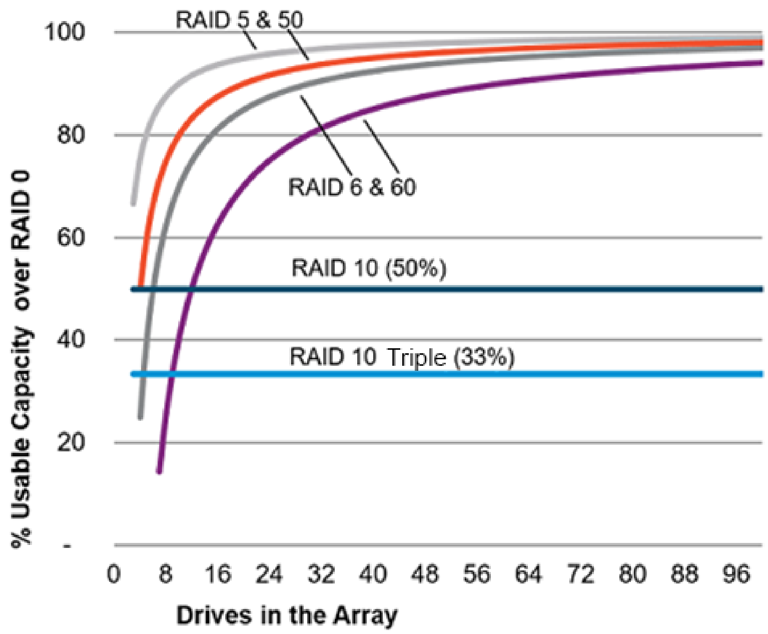


Figure 2: HP Storage Efficiency [3]

Figure 5.1 shows a general plot of the storage efficiency development over a increasing amount of disks. The plot assumes the group size 2 for RAID-50 and RAID-60. [3]

At last consider following suggestions by HP:

- RAID 1/10 Triple: Optimize for fault tolerance and write performance.
- RAID 6/60: Optimize for fault tolerance and usable capacity.
- RAID 1/10: Optimize for write performance.
- RAID 5/50: Optimize for usable capacity.

# Fun Facts

It appears to be, that RAID was an abreviation for "Redundant Array of Inexpensive Disks" befoire beeing modified to the more known version of: "Redundant Array of Independent Disks" Source is: trust me brother, cannot be bothered to look it up.

## List of Figures



## List of Tables



## References


[1] P. Chen, E. Lee, G. Gibson, R. Katz, and D. Patterson, "Raid: High-performance, reliable secondary storage," tech. rep., Carniege Mellon University, n.d.

[2] R. Alagappan, "Cs 537: Raid," tech. rep., University of Wisconsin, 2019.

[3] HPE, "Hpe smart array sr gen10 controller user guide," tech. rep., HP, 2024.

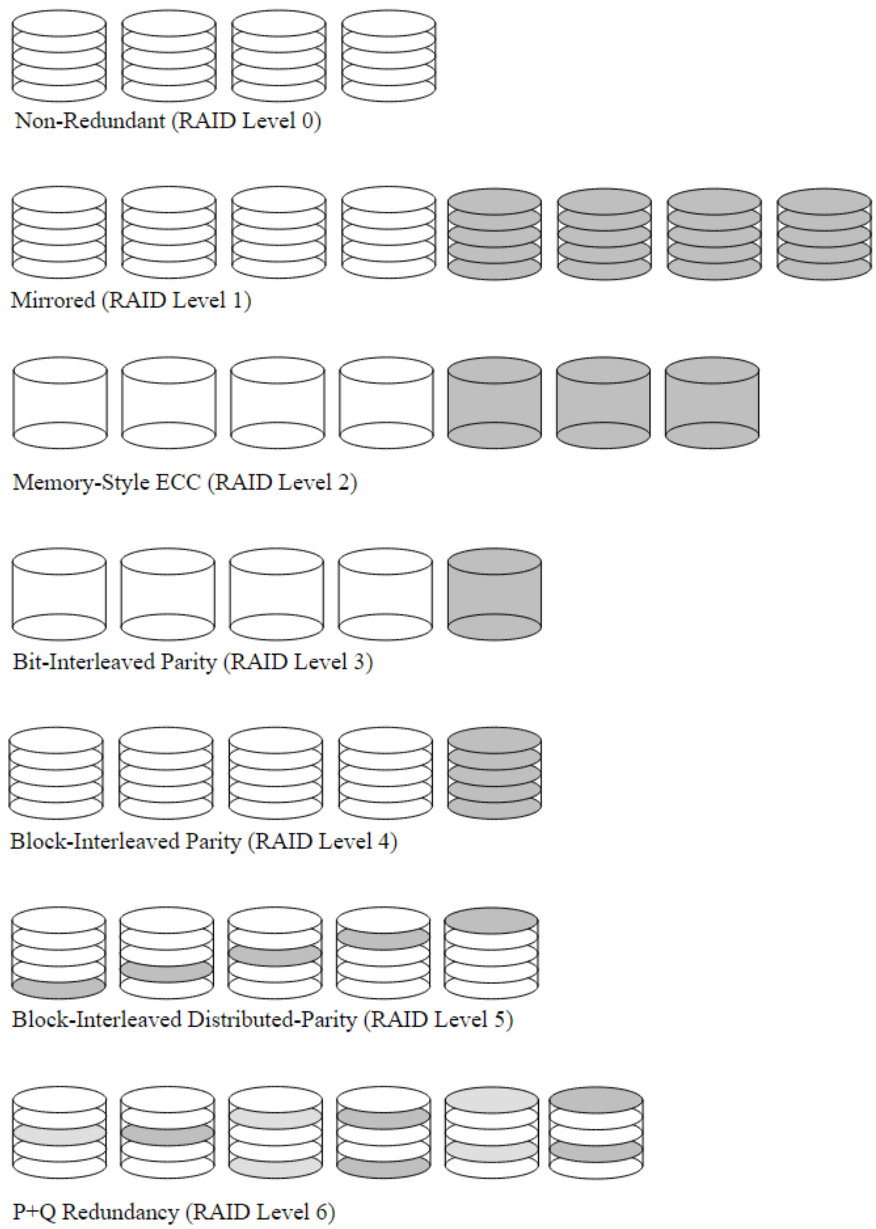


Figure 3: Raid Overview [1]