## ©ladaprec

## Adaptec ${ }^{\circledR}$ RAID 10 and its Alternatives

## Introduction

RAID 10 arrays are typically used in environments that require uncompromising availability coupled with exceptionally high throughput for the delivery of data located in secondary storage. In recent years a number of mutations of RAID 10 have been developed with similar capabilities. This paper presents one of the popular alternative implementations and discusses the relative advantages and disadvantages of RAID 10 and this alternative.

## RAID 10

A RAID 10 array is formed using a two-layer hierarchy of RAID types. At the lowest level of the hierarchy are a set of RAID 1 sub-arrays i.e., mirrored sets. These RAID 1 sub-arrays in turn are then striped to form a RAID 0 array at the upper level of the hierarchy. The collective result is a RAID 10 array. The figure below demonstrates a RAID 10 comprised of two RAID 1 sub-arrays at the lower level of the hierarchy. They are sub-arrays $A$ (comprised of disks $A_{1}$ and $A_{2}$ ) and $B$ (comprised of disks $B_{1}$ and $B_{2}$ ). These two sub-arrays in turn are striped using the strips $1_{A}$, $1_{B}, 2_{A}, 2_{B}, 3_{A}, 3_{B}, 4_{A}, 4_{B}$ to form a RAID 0 at the upper level of the hierarchy. The result is a RAID 10.


Figure 1 illustrates a RAID 10 array, with each disk in the array participating in exactly one mirrored set, thereby forcing the number of disks in the array to be even.

Let us now look at some of the salient properties of RAID 10. Consider a RAID 10 comprised of d disks and N mirrored sets (i.e., constituent RAID 1 sub-arrays). Since each disk in the array participates in exactly one mirrored set, $d=2 N$.
a) RAID 10 arrays do not require any parity calculation at any stage of their construction or operation.
b) RAID 10 arrays are generally deployed in environments that require a high degree of redundancy. The ability to survive multiple disks failures is a fundamental property of RAID 10. In fact the maximum number of disk failures a RAID 10 array can withstand is $\mathrm{d} / 2=\mathrm{N}$.

What about the number of combinations of failed disks that a RAID 10 array can sustain? The number of ways in which $k$ disks can fail is given by ${ }^{N} C_{k} \cdot 2^{k}$, since there are ${ }^{N} C_{k}$ ways in which to choose k mirror groups from N possible choices, and 2 ways in which to choose a disk within each mirror group. Therefore the total number of combinations of failed disks that a RAID 10 can support is:

$$
\begin{aligned}
& { }^{N} C_{1} \cdot 2^{1}+{ }^{N} C_{2} \cdot 2^{2}+\ldots+{ }^{N} C_{N} \cdot 2^{N} \\
& =(2+1)^{N}-1 \\
& =3^{N}-1
\end{aligned}
$$

Figure 1

Thus, for a 4 drive RAID 10 containing 2 mirrored sets, the number of combinations in which disks can fail without the array being rendered inoperable is $3^{2}-1=8$. In fact, these combinations may be enumerated as follows, with each possible set of failed disks listed within braces. They are: $\left\{A_{1}\right\},\left\{A_{2}\right\},\left\{B_{1}\right\},\left\{B_{2}\right\}$, $\left\{A_{1}, B_{1}\right\},\left\{A_{2}, B_{2}\right\},\left\{A_{1}, B_{2}\right\}$, and $\left\{A_{2}, B_{1}\right\}$.
c) RAID 10 ensures that if a disk in any constituent mirrored set fails, its contents can be extracted from the functioning disk in its mirrored set. Thus, when a RAID 10 array has suffered the maximum number of disk failures it is capable of withstanding, its throughput rate is no worse than that of a RAID 0 with $N$ disks. In fact, any combination of N contiguous independent strips can be read concurrently. The term "independent strip" is used to denote a strip in a collection of strips that is not a mirror of any other strip within that collection.
d) A RAID 10 array that is in a nominal state can improve the throughput of read operations by allowing concurrent reads to be performed on multiple disks in the array. For example, if the strips $1_{A}, 1_{B}, 2_{A}, 2_{B}$ are to be read from the array given in Figure 1, it is clear that all four strips can be read concurrently from the disks $\mathrm{A}_{1}, \mathrm{~B}_{1}$, $A_{2}$ and $B_{2}$ respectively.

RAID 1E
While RAID 10 has been traditionally implemented using an even number of disks, some hybrids can use an odd number of disks as well. Figure 2 illustrates an example of a hybrid RAID 10 array comprised of five disks; A, B, C, D and E. In this configuration, each strip is mirrored on an adjacent disk with wrap-around. In fact this scheme - or a slightly modified version of it - is often referred to as RAID 1 E and was originally proposed by IBM. Let us now investigate the properties of this scheme.

When the number of disks comprising a RAID 1 E is even, the striping pattern is identical to that of a traditional RAID 10 , with each disk being mirrored by exactly one other unique disk. Therefore, all the characteristics for a traditional RAID 10 apply to a RAID 1 E when the latter has an even number of disks. However, RAID 1E has some interesting properties when the number of disks is odd.
a) Just as in the case of traditional RAID 10, RAID 1E does not require any parity calculation either. So in this category, RAID 10 and RAID 1E are equivalent.
b) The maximum number of disk failures a RAID $1 E$ array using d disks can withstand is $\lfloor\mathrm{d} / 2 \mathrm{~L}$. When $d$ is odd, this yields a value that is the equal to that of a traditional RAID 10 while utilizing one additional disk. What about the number of combinations of disk failures that RAID 1E can support? It turns out that RAID 1E is very peculiar in this characteristic when d is odd.


Figure 2

Assume for the sake of notational convenience that $\lfloor d / 2\rfloor=p$. Then the number of ways in which $k$ disks can fail is $d \cdot{ }^{p-1} C_{k-1}$, since there are d ways to choose the first disk and ${ }^{p-1} C_{k-1}$ ways to choose the remaining $k-1$ disks from p-1 possible choices.

Therefore, the total number of combinations of failed disks that this scheme can support is:

$$
\begin{aligned}
& d \cdot p-1 C_{0}+d \cdot d \cdot p-1 C_{1}+\ldots+d \cdot d \cdot 1 C_{p-1} \\
& =d \cdot\left({ }^{p-1} C_{0}+{ }^{p-1} C_{1}+\ldots+{ }^{p-1} C_{p-1}\right) \\
& =d \cdot 2^{p-1}
\end{aligned}
$$

Thus, for a 5 drive RAID 1E, the total number of combinations in which disks can fail without the array being rendered inoperable is $5 \cdot 2^{2-1}=10$. However, this result also indicates that as the value of $d$ increases, the ratio of the number of combinations of disk failures supported by RAID 1E using d disks decreases with respect to conventional RAID 10 using d-1 disks. In fact, for $d>9$, RAID 1E yields a lesser number of combinations! For instance, while a conventional RAID 10 using 10 disks can support $3^{5}-1=242$ combinations of disk failures, RAID 1E using 11 disks can support only $11 \cdot 2^{5-1}=176$ combinations. Clearly, RAID 10 is a superior choice when tolerance to a larger number of combinations of disk failures is considered important. An even more significant implication of this result is the following. Since a RAID 1E with an even number of disks is identical to a traditional RAID 10, a RAID 1E with 10 disks can support more combinations of failures than a RAID 1E with 11 disks. In general, a RAID 1E with 2 N disks can support more combinations of failures that a RAID 1E with $2 \mathrm{~N}+1$ disks, when $N \geq 5$. In other words, it is always preferable to utilize an even number of disks for a RAID 1E than an odd number if you desire a higher tolerance to disk failures. In other words, it is always preferable to use a traditional RAID 10!
c) When a RAID 1E array suffers the maximum number of disk failures it is capable of withstanding, i.e., $\lfloor d / 2\rfloor$, the number of contiguous independent strips that can be accessed concurrently can be less than $\lceil\mathrm{d} / 27$. In fact, the minimum number of independent continguous strips that can be accessed is \d/2」. For example, consider the RAID 1E array displayed in Figure 2. Assume that disks A and C have failed. In this scenario, it is clear that the contiguous strips 4, 5 and 6 cannot be read concurrently although three disks remain operational. Thus the throughput of a RAID 1E with d disks - where $d$ is odd - may be no higher under specific access patterns than that of a RAID 10 with $\mathrm{d}-1$ disks when both arrays experience the maximum number of sustainable disk failures.
d) Just as in the case of a traditional RAID 10 implementation, RAID 1E in a nominal state can improve the throughput of read operations by allowing concurrent reads to be performed on multiple disks in the array. The fact that there are more disks than there are mirror sets should intuitively suggest as much.

## Conclusion

RAID 1E offers a little more flexibility in choosing the number of disks that can be used to constitute an array. The number can be even or odd. However, RAID 10 is far more robust in terms of the number of combinations of disk failures it can sustain even when using lesser number of disks. Furthermore, a RAID 10 guarantees a throughput rate that is always equal to that which is obtainable from the concurrent use of all its functioning disks. In contrast, specific access patterns may not lend themselves to the concurrent use of all functioning disks under RAID 1E. Therefore, if extremely high availability and throughput are of paramount importance to your applications, RAID 10 should be the configuration of choice!

## ©

