

## Why RAID Cannot Be Considered a Storage Security Solution

In the storage realm RAID architecture continues to be very popular and is widely used by different vendors because it allows for the combination of different hard drives into one fast, reliable and spacious storage device that satisfies nearly all enterprise data storage needs. However, along with all the well known benefits of RAID architecture a common misconception continues to exist; many IT professionals still believe that the data stored on RAID devices is secure. This false belief stems from the basic concept of RAID – distributing the data among many hard drives which disrupts files' formats and makes the stealing of one particular hard drive from the RAID system useless for insider. This paper highlights the threats to data in a RAID architecture and outlines why additional data protection procedures should be employed to ensure complete protection and compliance with the ever-increasing regulatory mandates for securing sensitive data.

There are three key concepts in a RAID architecture: mirroring, the copying of data to more than one disk; striping, the splitting of data across more than one disk; and error correction, where redundant data is stored allowing problems to be detected and possibly fixed (known as fault tolerance). [1]

### Mirroring

If only mirroring is implemented, no data is distributed among hard drives at all and there are no obstacles for a thief – the same data is stored on more than one hard drive and the loss of one hard drive means compromising all data within the RAID. Fortunately RAID mirroring is rarely implemented as a standalone solution.

### Striping

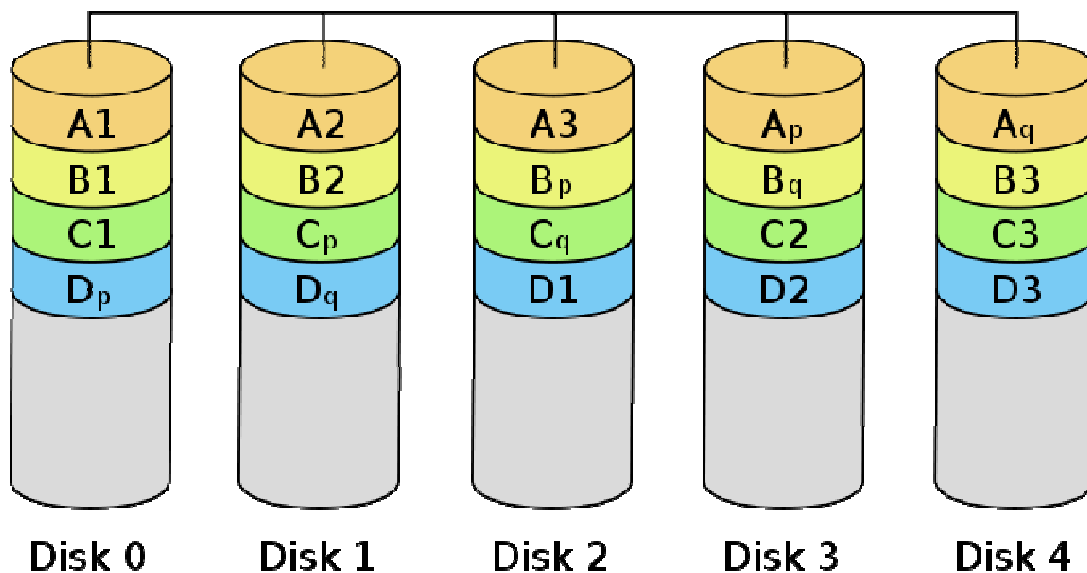
Striping means that data block is split into several smaller pieces before being written to the device (the number of pieces depends on the number of hard drives) and each piece is written to a separate hard drive. The size of the piece varies at different RAID levels: RAID2 stripes the data at the bit level so the size of the piece is one bit; RAID3 – at the byte level; RAID4, RAID5 and RAID6 – on the block level. [2] The size of one block on all modern hard drives is 512 bytes. We will analyze only block-level stripes because they are used in most popular storage solutions. For example, NetApp's FAS6000 Series storage employs RAID6 for higher data availability with little or no performance loss. [3]

Let's imagine that a 1 MB text file with credit card data called `cc_data.txt` is written to a RAID6 device consisting of 5 hard drives. The file `cc_data.txt` contains text records with cardholder name, card number, expiry date and security code delimited by tabs:

<i>cc_data.txt</i>				
LAKSHMI MANTRA	4521 3488 3211 3453	09/11	789	
ALEXEY RAEVSKY	4276 3800 4567 1234	08/10	365	
JOHN DOE	4562 2345 6789 3478	09/12	123	
WU MINGSHI	4323 8900 3255 9877	01/10	354	
RICHARD ROE	4134 4523 7845 9009	07/09	987	
...				

Let the size of one record in the file be 64 bytes, so the whole file contains 16,384 records. In RAID6, the data is striped at the block level with two blocks of parity information distributed across all member disks.

# RAID 6



RAID 6 Data Layout

On the picture A1, A2 and A3 signify data blocks, A<sub>p</sub> and A<sub>q</sub> – parity information and so forth.

The file from our example will be split into 512-bytes pieces (2048 pieces total) that are written to RAID in the following order: first three pieces to the drives 0, 1 and 2, fourth and fifth – to the drives 0 and 1, sixth – to the drive 4 and so on as displayed in the picture above. Each drive now contains three-fifths of the original file. If even a single hard drive is stolen from this RAID array, the thief has obtained 9830 credit card records. It is not the loss of the whole file, but the damage from the loss of and ultimately illegal use of these records is significant for the company.

According to Ponemon Institute research, the average cost of a data breach for a company is \$202 per record of stolen data so in our example the cost of losing one hard drive from the RAID6 will be nearly \$2M.

Of course text files are not often used in today's business environments, but even with other types of files including database tables, the situation will be similar to the example outlined above. In nearly every case, an insider thief will already know the file format making stealing the data is only slightly more complicated than with text file, but this is a matter of file complexity only.

In general, we can say that if the data record size is smaller than a RAID stripe piece size, the loss of even a single hard drive from this RAID carries a significant risk of compromising the data, which must be considered.

### **Error correction**

As seen from the previous examples, storing redundant data across the RAID does not limit the ability of a thief to recover data from a stolen hard drive. The exception to this is a situation where error-correction data is stored on a separate hard drive and the thief stole this specific hard drive. In this case, with no useful information on the stolen drive, the company is very lucky. However, hoping for the best and relying on fortune or fate is no substitute for proper network and [data storage security](#) planning and procedures.

### **Summary**

We have seen that mirroring does not hide data at all. And though striping data in RAID devices at the block level may distribute files across the RAID disks, it still leaves data accessible to a thief who can steal one hard drive from the RAID array. Significant amounts of data will be compromised and additional protection techniques, such as [server data encryption](#), should be considered to eliminate the financial, regulatory and reputation risks.

### **References**

1. RAID. <http://en.wikipedia.org/wiki/RAID>
2. Standard RAID levels. [http://en.wikipedia.org/wiki/Standard\\_RAID\\_levels](http://en.wikipedia.org/wiki/Standard_RAID_levels)
3. NetApp FAS6000 Series. <http://media.netapp.com/documents/fas6000.pdf>