RAID (Redundant array of independent disks)
World of Computer Science. 2007.

Full Text:

RAID is an acronym for Redundant Array of Independent Disks. It is a method of spreading stored information across several hard drives. Information can be duplicated on several disks, hence the term **redundancy**. By spreading **data** across several hard disks, users can get both faster performance and greater **data security**. Depending upon the RAID system being used, a failed drive can be replaced without having to shut down the **host system**.

In 1987, a group at University of California, Berkley, published a paper proposing RAID. The technology was designed to increase the speed of secondary **memory**, so as to keep pace with the increasing speed of central processing units.

There are several levels of RAID configurations. RAID 0 writes data across the different drives, one segment of a drive at a time. Also called striping, it is a linear means of writing the data, where writing begins at a segment of the first drive and then works progressively through the drives.

RAID 1 writes data to two drives at the same time. If one drive fails, the data is not lost, since it resides on the other drive. This process is called mirroring, and it provides total redundancy. Besides providing the protection associated with two copies of the data, mirroring speeds up the reading of data. Different blocks of data can be read simultaneously from the two drives, and then compiled together in the original order. There are several disadvantages to RAID 1. Redundancy is inefficient with respect to space, and the need for more data storage space affects the cost. Also, RAID 1 is often implemented in software rather than **hardware** (also called host based RAID). Software performance demands much from the **CPU**, and so can lower the overall performance of the system. The reliability it offers can be worth the expense, however, in **applications** where data loss could be disastrous, such as accounting or payroll settings.

The original list of RAID configurations included RAID 2. Instead of creating a backup of all the data, RAID 2 uses a so-called error correction **code**. The code detects and corrects single errors during their input, so that mirroring of data is not necessary. The storage requirements for the code are considerable, however. Due to the inefficiency for small data transfers and the high cost of the additional drives required for the code, RAID 2 has not been commercially available.

RAID 3, also called RAID 0+1, is a combination of striping and **parity** checking. Parity checking involves determining whether each given block of each drive has an odd or even **value**. When all the values across a stripe are added up, the parity value is obtained. Once a drive's parity value is known, then the failed area of a failed drive can be determined and the information rebuilt on another drive. Parity checking provides redundancy without having to double drive capacity. Because of its efficiency in large data transfers, RAID 3 is used where large files are being stored and where quickness is necessary. Video production and live streaming of images utilize RAID 3 technology.

RAID 4 is similar to RAID 3 except that the size of the stripe written across the drives is larger. This means that the information is spread out across the drives more, and that each drive will not contain exactly the same information. The drives are more independent of one another. The parity information about all the drives is stored only on one of the drives. While this spreads the writing of the information to the drives, recovery of data in the event of a drive failure is more complex and time consuming.

Finally, RAID 5 is a combination of striping and a form of parity checking where the parity information is distributed across all the drives. This improves the speed at which data can be read and makes for

the most efficient storage of data. Rebuilding a failed drive can still be onerous. Nonetheless, RAID 5 is widely used in servers for **databases**, **e-mail**, newsgroups, and **intranet**s.

The most commonly used RAID implementations are RAIDs 1, 3, and 5. Each has their niche with respect to application. Technological advances have created newer forms of RAID (6, 7, 10, 53). Each offers variations on the partitioning of data based on striping or mirroring. Some of the improvements in speed and security of data storage come with greatly increased price tags. The choice of a RAID system becomes one of a balanced consideration of the data storage needs and budget.

Source Citation (MLA 8th Edition)
"RAID (Redundant array of independent disks)." *World of Computer Science*, Gale, 2007. *Science In Context*, http://link.galegroup.com/apps/doc/CV2424500495/SCIC?u=txshracd2598&sid=SCIC&xid=dd89962e. Accessed 30 Apr. 2018.

Gale Document Number: GALE|CV2424500495