RAID 5

**Definition**

RAID 5 is a [redundant array of independent disks](https://searchstorage.techtarget.com/definition/RAID) configuration that uses [disk striping](https://searchstorage.techtarget.com/definition/RAID-0-disk-striping) with [parity](https://searchstorage.techtarget.com/definition/parity). Because data and parity are striped evenly across all of the disks, no single disk is a bottleneck. Striping also allows users to reconstruct data in case of a disk failure

RAID 5 evenly balances reads and writes, and is currently one of the most commonly used RAID methods. It has more usable storage than RAID 1 and RAID 10 configurations, and provides performance equivalent to RAID 0.

RAID 5 groups have a **minimum of three hard disk drives** (HDDs) and no maximum. Because the parity data is spread across all drives, RAID 5 is considered one of the most secure RAID configurations.

Space efficiency :- 1-1/n

Fault tolerance :-one drive failure

**Working of RAID 5**

The benefits of RAID 5 primarily come from its combined use of disk striping and parity.

1. **Striping** is the process of storing consecutive segments of data across different storage devices, and allows for better throughput and performance. Disk striping alone does not make an [array](https://searchstorage.techtarget.com/definition/array) fault tolerant, however. Disk striping combined with parity provides RAID 5 with redundancy and reliability.
2. RAID 5 used **parity** instead of mirroring for data redundancy. When data is written to a RAID 5 drive, the system calculates parity and writes that parity into the drive. While mirroring maintains multiple copies of data in each volume to use in case of failure, RAID 5 can rebuild a failed drive using the parity data, which is not kept on a fixed single drive

By keeping data on each drive, any two drives can combine to equal the data stored on the third drive, keeping data secure in case of a single drive failure. Drives can be [hot swapped](https://whatis.techtarget.com/definition/hot-swap) in RAID 5, which means a failed HDD can be removed and replaced without downtime

**Logic:-**

First we need to remind you XOR definition:

XOR function result is equal 1 if both arguments are differe

XOR (1, 0) = 1

XOR (0, 1) = 1

XOR function output is equal 0 if both arguments are same.

XOR (1, 1) = 0

XOR (0, 0) = 0

Now let us assume we have 3 drives with the following bits:

| 101 | 010 | 011 |

And we calculate XOR of those data and place it on 4th drive

XOR (101, 010, 011) = 100     (XOR (101,010) = 111 and then XOR (111, 011) = 100

So the data on the four drives looks like this below:

| 101 | 010 | 011 | **100** |

Let’s assume the second drive has failed. When we calculate XOR all the remaining data will be present from the missing drive.

| 101 | **~~010~~**| 011 | 100 |

XOR (101, 011, 100) = 010

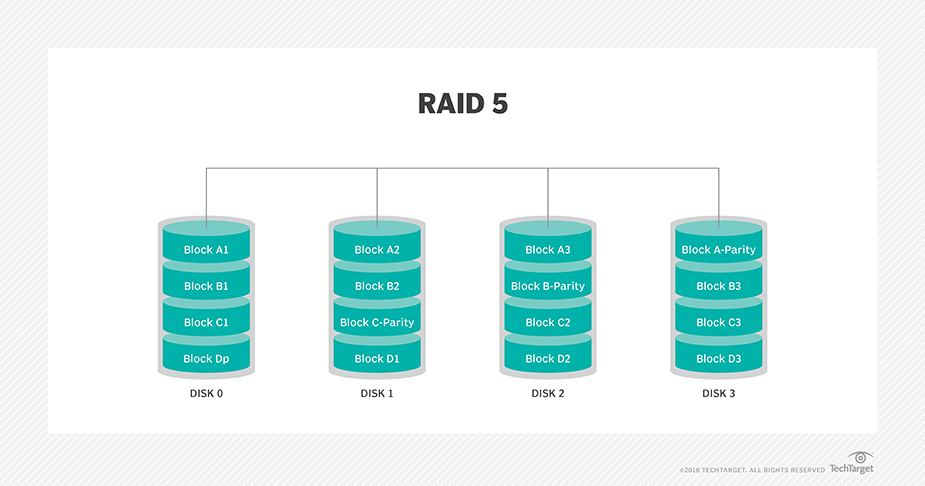
You can check the missing other drives and XOR of the remaining data will always give you exactly the data of your missing drive.| 101 | 010 | **~~011~~** | 100 |

XOR (101, 010, 100) = 011

What works for 3 bits and 4 drives only, works for any number of bits and any number of drives. Real RAID 5 has the most common stripe size of 64k (65536 \* 8 = 524288 bits )

So the real XOR engine only needs to deal with 524288 bits and not 3 bits as in our exercise. This is why the RAID 5 needs a very efficient XOR engine in order to calculate it fast.

So when adding one drive for parity you will be able to rebuild the missing data in case of any drive failure.



**Adavantages**

1. Fast, reliable reads are major benefits.
2. This RAID configuration also offers inexpensive data redundancy and fault tolerance.
3. Data can be accessed and read even while a failed drive is being rebuilt. When drives fail, the RAID 5 system can read the information contained on the other drives and recreate that data, tolerating a single drive failure.

**Disadvantages**

1. Writes tend to be slower, because of the parity data calculation.
2. Longer rebuild times are one of the major drawbacks of RAID 5, and this delay could result in data loss.
3. Because of its complexity, RAID 5 rebuilds can take a day or longer, depending on controller speed and work load.
4. If another disk fails during the rebuild, then data is lost forever.