**RAID** (originally **R**edundant **A**rray of **I**nexpensive **D**isks; now commonly **R**edundant **A**rray of **I**ndependent **D**isks) is a data storage virtualization technology that combines multiple disk drive components into a logical unit for the purposes of data redundancy and performance improvement

The term "**RAID**" was invented by David Patterson, Garth A. Gibson, and Randy Katz at the University of California, Berkeley in 1987, in a paper titled "A Case for Redundant Arrays of Inexpensive Disks (**RAID**)" in June 1988 at the SIGMOD conference

**RAID 0** – 2 drives working for redundancy

* Backup
* **RAID 0** comprises striping (but no parity or mirroring). This level provides no data redundancy nor fault tolerance, but improves performance through parallelism of read and write operations across multiple drives. RAID 0 has no error detection mechanism, so the failure of one disk causes the loss of all data on the array.

**RAID 1** – Slower, contains 2 disks

* **RAID 1** comprises mirroring (without parity or striping). Data are written identically to two (or more) drives, thereby producing a "mirrored set". The read request is serviced by any of the drives containing the requested data. This can improve performance if data is read from the disk with the least seek latency and rotational latency. Conversely, write performance can be degraded because all drives must be updated; thus the write performance is determined by the slowest drive. The array continues to operate as long as at least one drive is functioning

**RAID 2** – 2 drives contain the exact same files

* Hamming code for error correction
* Bits of data vs. blocks
* **RAID 2** comprises bit-level striping with dedicated Hamming-code parity. All disk spindle rotation is synchronized and data is striped such that each sequential bit is on a different drive. Hamming-code parity is calculated across corresponding bits and stored on at least one parity drive. This level is of historical significance only. Although it was used on some early machines (e.g. the Thinking Machines CM-2), it is only recently used by high-performance commercially available systems

**RAID 3** – Uses byte-level striping with a dedicated parity disk

* Cannot service multiple request simultaneously
* Suitable for application that demand highest transfer rates with long sequential reads & writes.
* **RAID 3** comprises byte-level striping with dedicated parity. All disk spindle rotation is synchronized and data is striped such that each sequential byte is on a different drive. Parity is calculated across corresponding bytes and stored on a dedicated parity drive. Although implementations exist, RAID 3 is not commonly used in practice

**RAID 4** – Uses block-level striping with dedicated parity disk vs. a byte

* Very uncommon, NETApp, enterprise level company
  + WAFL – write anywhere file layout
* Good for Metadata
* **RAID 4** comprises block-level striping with dedicated parity. This level was previously used by NetApp, but has now been largely replaced by a proprietary implementation of RAID 4 with two parity disks, called RAID-DP

**RAID 5** – Block-level striping with distributed parity

* Requires all drive but 1 be present
  + Lost Drive does not mean lost data
* **RAID 5** comprises block-level striping with distributed parity. Unlike in RAID 4, parity information is distributed among the drives. It requires that all drives but one be present to operate. Upon failure of a single drive, subsequent reads can be calculated from the distributed parity such that no data is lost. RAID 5 requires at least three disks. RAID 5 is seriously affected by the general trends regarding array rebuild time and chance of failure during rebuild. In August 2012, Dell posted an advisory against the use of RAID 5 in any configuration and of RAID 50 with "Class 2 7200 RPM drives of 1 TB and higher capacity"

**RAID 6** – Block-level striping with distributed parity

* Requires all drive but 1 be present
  + Lost Drive does not mean lost data
* Performance/Speed
* **RAID 6** comprises block-level striping with double distributed parity. Double parity provides fault tolerance up to two failed drives. This makes larger RAID groups more practical, especially for high-availability systems, as large-capacity drives take longer to restore. As with RAID 5, a single drive failure results in reduced performance of the entire array until the failed drive has been replaced. With a RAID 6 array, using drives from multiple sources and manufacturers, it is possible to mitigate most of the problems associated with RAID 5. The larger the drive capacities and the larger the array size, the more important it becomes to choose RAID 6 instead of RAID 5. RAID 10 also minimizes these problems.

**RAID** devices:

* Use many different architectures to achieve desired balance, performance and fault tolerance.
* The architectures are called levels.
* Standard RAID levels 0-6 & 10
  + Level 10 (1+0) is the most common + mirror stripes
  + RAID 0+1: creates a second striped set to mirror a primary striped set. The array continues to operate with one or more drives failed in the same mirror set, but if drives fail on both sides of the mirror the data on the RAID system is lost
  + RAID 1+0: creates a striped set from a series of mirrored drives. The array can sustain multiple drive losses so long as no mirror loses all its drives
* Some devices use more than 1 level called hybrid or nested
* Some vendors offer non-standard proprietary RAID levels (Nested/hybrid RAID).