Understanding RAID

RAID Defined

The acronym "RAID" stands for Redundant Array of Inexpensive (or now more commonly known as "Independent") Disks. Over the past several years RAID has grown increasingly popular because of its I/O performance-enhancing and disk fault-tolerant features. Today RAID is used all over the world in entry-level to enterprise-wide computing environments.
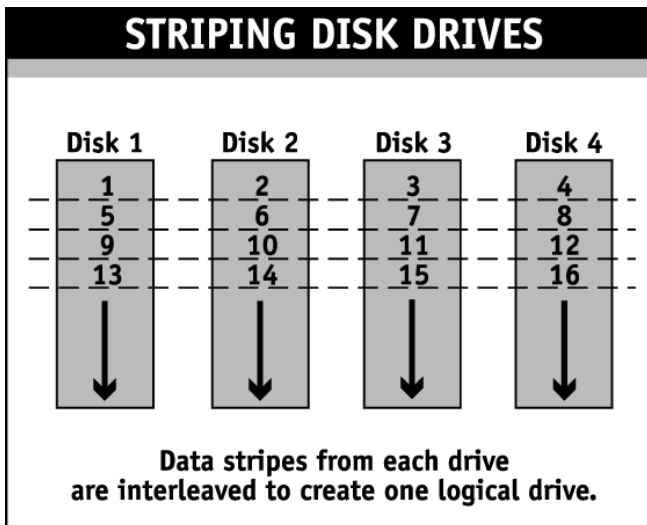
The Genesis of RAID

In 1987, Patterson, Gibson, and Katz of the University of California, Berkeley, published a paper entitled <u>A Case for Redundant Arrays of Inexpensive Disks (RAID)</u>[1]. This paper described various types of disk arrays, referred to by the acronym RAID. The purpose of RAID was to combine multiple small, inexpensive disk drives into a single logical drive or "disk array" that yields performance exceeding that of a more expensive high speed disk drive. In addition to improving performance, disk arrays can also provide disk fault-tolerance by redundantly storing information in various ways.

Five types of disk array architectures, RAID-1 through RAID-5, were defined by the Berkeley paper. Each of these RAID levels provides disk fault-tolerance and each offer different trade-offs in features and performance. In addition to these five redundant disk array architectures, it has become popular to refer to a non-redundant array of disk drives as a RAID-0 array.

The Concept of Striping

The precursors to today's RAID arrays were groups of striped drives, now referred to as RAID-0 arrays. A striped array is created by logically combining two or more disk drives into a single logical storage unit. However, instead of concatenating the storage space of each drive end-to-end, the logical space is organized by partitioning each drive into stripes, which may be as small as one sector (512 bytes) or as large as several megabytes. These stripes are then interleaved round-robin, so that the combined space is composed alternately of stripes from each drive. The type of application environment, random I/O-intensive or throughput-intensive, determines whether large or small stripes, should be used for best performance.

I/O-intensive environments like database, transaction processing, and general multi-user office applications, access many small random records. In these environments, performance is optimized by creating stripes large enough so that each record resides entirely on a single drive. Since each I/O will access only one physical drive of the array, I/O operations can be evenly distributed across the

array, allowing each drive to work on a different I/O operation. This maximizes the number of simultaneous I/O operations that can be performed by the array.
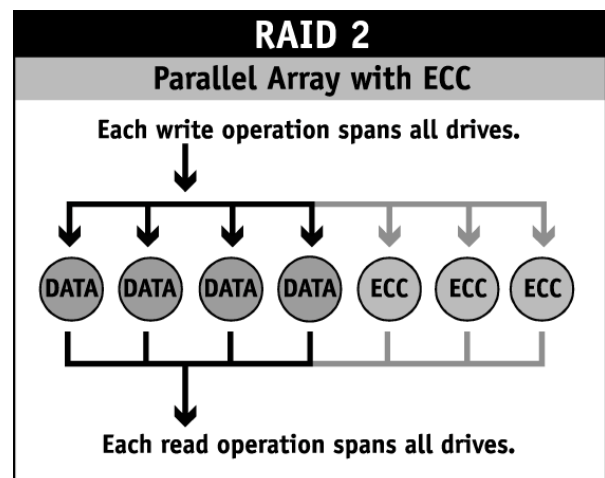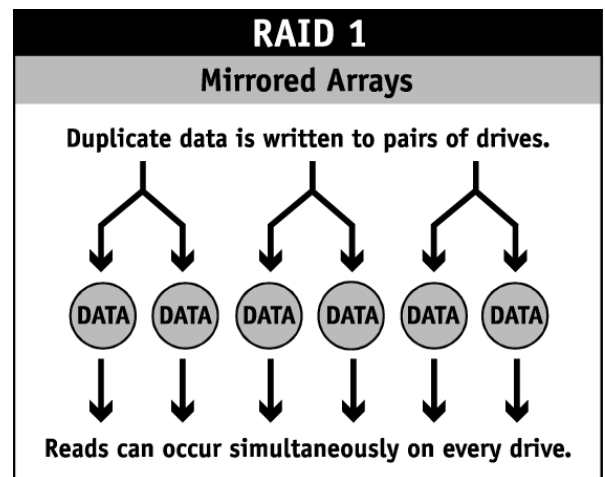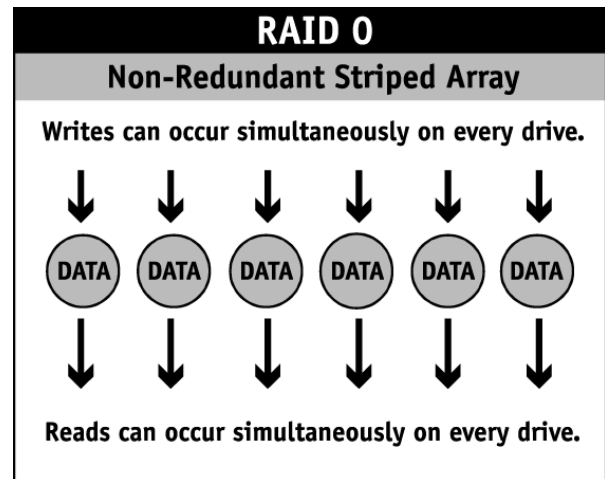
Throughput-intensive environments like video-on-demand, image editing, pre-press, and internet data download, access large sequential records. Using small stripes relative to the record size results in each record spanning across all the drives in the array. This allows large records to be accessed faster since each record is broken up and transferred in parallel across multiple drives.

RAID Levels Defined

RAID-0 arrays are groups of striped disk drives with no data redundancy and thus no fault tolerance. RAID-0 arrays can be configured with large stripes for I/O-intensive applications, or small stripes for throughput-intensive applications. Since RAID-0 does not provide redundancy, a single drive failure will cause the array to crash. However, RAID-0 arrays deliver the best performance and data storage efficiency of any array type.

RAID-1, also known as disk mirroring, consists of pairs of disk drives that store duplicate data, yet appear to the computer as a single drive. If one drive fails, the other drive of the pair is still available. A pair of mirrored drives has better read throughput than an individual drive because both drives of the pair can perform reads simultaneously. However, write throughput is the same as for a single drive since every write must go to both drives of the pair. Striping is not used with a single pair of drives. However, multiple pairs may be striped together to appear as a single larger array. This configuration is sometimes referred to RAID 0+1 or RAID-10. RAID-1 offers good performance and fault tolerance but has the least storage efficiency of any RAID level.

RAID-2 arrays are striped similar to RAID-0 except that disk fault tolerance is achieved by devoting some drives to storing ECC information. Since all disk drives today embed ECC information within each sector, RAID-2 is not used.

RAID 0
Non-Redundant Striped Array
Writes can occur simultaneously on every drive.
Reads can occur simultaneously on every drive.

RAID 1
Mirrored Arrays
Duplicate data is written to pairs of drives.
Reads can occur simultaneously on every drive.

RAID 2
Parallel Array with ECC
Each write operation spans all drives.
Each read operation spans all drives.

RAID-3, is similar to RAID-2 except that a single drive is devoted to storing parity information instead of one or more drives storing ECC information. If a drive fails, any missing stripe can be recovered by calculating the exclusive OR of similarly positioned stripes on the remaining drives. RAID-3 requires that records span across all drives in the array, so that pieces of each record are transferred in parallel, maximizing transfer rate. Therefore, the stripe size for RAID-3 must be small relative to the record size. As a result, RAID-3 can perform only one I/O at a time, limiting its use to single-user systems.

RAID-4 is identical to RAID-3 except that the stripes are larger than the typical record. As a result, records typically reside entirely on a single drive in the array. This allows multiple simultaneous read operations and therefore, greater throughput in multi-tasking and multi-user systems. However, since all write operations must update the single parity drive, only one write can occur at a time. This architecture offers no significant advantages over RAID-5 and its write performance is slower.

RAID-5 avoids the write bottleneck caused by the single dedicated parity drive of RAID-4. RAID-5 uses "rotating parity", evenly distributing parity information among all drives in the array. Multiple write operations can be processed simultaneously, resulting in improved throughput over RAID-4. Like RAID-3 and RAID-4, the equivalent of one drive's capacity is sacrificed for the array's parity data. RAID-5 arrays are versatile because they can be configured with small stripes to yield performance characteristics similar to RAID-3 or with large stripes for multi-tasking and multi-user environments. Multiple RAID-5 parity groups may be striped together to appear as a single larger array. This configuration is sometimes referred to RAID 0+5 or RAID-50.
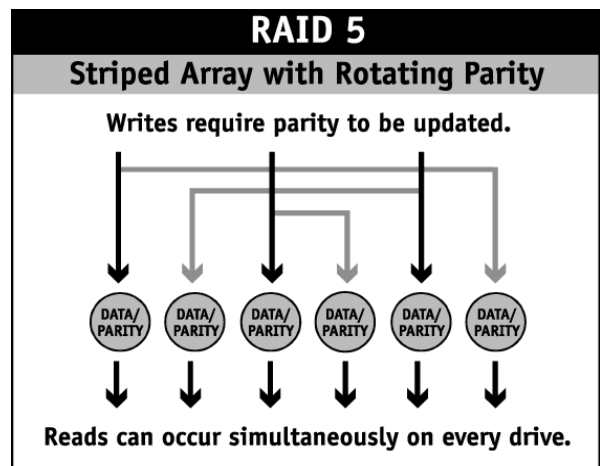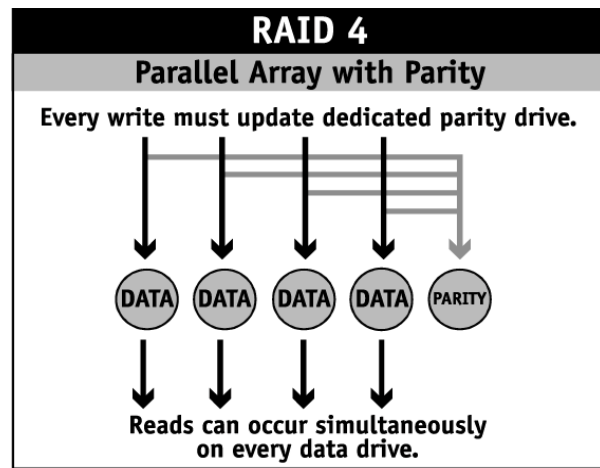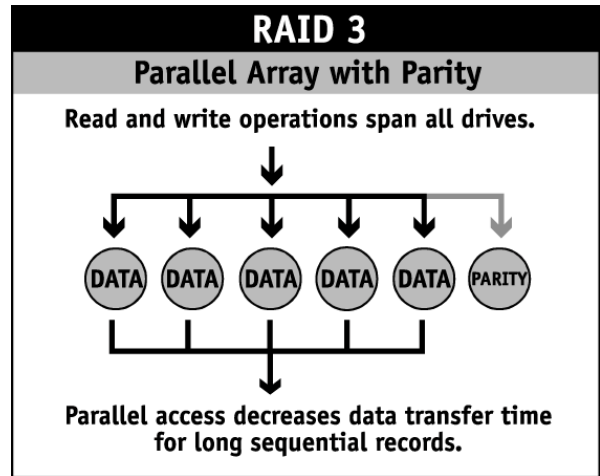
RAID Level Feature Summary							
RAID Level	Disk Fault Tolerance	Efficient Storage	Good for Multi-user	Good for Data Streaming	Good Read Performance	Good Write Performance	Commonly Used Today
RAID-0		✓	✓	✓	✓	✓	✓
RAID-1	✓		✓	✓	✓	✓	✓
RAID-2	✓	✓		✓	✓		
RAID-3	✓	✓		✓	✓		
RAID-4	✓	✓		✓	✓		
RAID-5	✓	✓	✓	✓	✓		✓

Note: Performance increases as more drives are included in an array.

[1] D. A. Patterson, G. Gibson, and R. H. Katz "A Case for Redundant Arrays of Inexpensive Disks (RAID)", Report No. UCB/CSD 87/391, University of California, Berkeley, CA 1987.