**EXPERIMENT NO.2**

**Aim:**To study various RAID levels in storage.

**Thoery:**

**RAID**-Redundant Array of Inexpensive Disks - is a method of combiningseveral hard drives into one logical unit. It can offer fault toleranceand higher throughput levels than a single hard drive or group ofindependent hard drives.RAID is a mature technology that speeds up data access while at thesame time protecting your data from hard disk failure. RAID is quicklybecoming a necessary component in every network since data lossand downtime can prove both fatal and financially destructive. Mostnetworks are designed to provide instant access to massive amountsof data. More and more employees have to access customer and otherdatabases. Intranets and corporate Web sites provide access to hugedatabases online.

**Raid Components and Concepts:**

**Logical Arrays** as a split or combination ofPhysical Arrays, which in turn are one or more Physical Drives thatare simply the individual hard disks that comprise these arrays.

**Mirroring** refers to complete redundancy of data on identical disks.The data that is being written on one Logical Array is completelyduplicated on a similar array thereby providing 100% dataredundancy. The cost associated with mirroring is that the amount ofavailable storage is reduced by 50%; writes are slightly slower albeitreads are faster in some situations.

**Striping** refers to a technique that allows Physical Drives in a LogicalArray to be used in parallel in order to gain in performance. In thistechnique, data is broken down in Byte or Block levels or stripes,where every Byte or Block is written to a separate disk in the array.Byte level can at times be a 512-byte sector, while Block size can beselected from variety of choices. The gain in performance is similarbetween Reads and Writes.

**Benefits of RAID:**

RAID provides increased storage capacities, and protects yourimportant data from hard drive failure.There are multiple benefits of using RAID:

* Reliability and Scalability
* Real-time data recovery with uninterrupted access when a harddrive fails
* System uptime and network availability and protection from loss
* Protection against data loss
* Multiple drives working together increase system performance

**RAID Levels:**

There are various RAID levels which are as follows:

**RAID-0 [Striping]:**

RAID 0 (also known as a stripe set or striped volume) splits ("stripes") data evenly across two or more disks, without parity information, redundancy, or fault tolerance. Since RAID 0 provides no fault tolerance or redundancy, the failure of one drive will cause the entire array to fail; as a result of having data striped across all disks, the failure will result in total data loss. This configuration is typically implemented having speed as the intended goal. RAID 0 is normally used to increase performance, although it can also be used as a way to create a large logical volume out of two or more physical disks.

**Minimum number of drives required: 2**

**Recommended Applications:**

Video Production and Editing

Image Editing

Pre-Press Applications

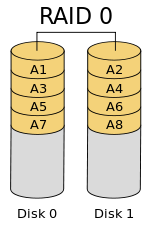
Any application requiring high bandwidth

**Advantages of RAID-0**

* High performance
* Very simple design, easy to implement
* No parity overhead
* No capacity loss because all storage is usable

**Disadvantages of RAID-0**

* Lack of fault-tolerance
* Failure of a single drive will result in loss of all data on the array
* Should never be used in mission critical environments



**RAID-1 [Mirroring]:**

RAID 1 consists of an exact copy (or mirror) of a set of data on two or more disks; a classic RAID 1 mirrored pair contains two disks. This configuration offers no parity, striping, or spanning of disk space across multiple disks, since the data is mirrored on all disks belonging to the array, and the array can only be as big as the smallest member disk. This layout is useful when read performance or reliability is more important than write performance or the resulting data storage capacity.

**Minimum number of drives required: 2**

**Recommended Applications:**

* Accounting, Payroll, and Financial
* Any application requiring very highavailability

**Advantages of RAID-1:**

* One Write or two Read possible operations permirrored pair
* Twice the Read transaction rate of singledisks, same Write transaction rate as single disks
* Fault tolerant
* Transfer rate per block is equal to that ofa single disk
* Easy to recover data in case of drivefailure, as no rebuild is necessary in caseof a disk failure, just a copy to thereplacement disk
* Easyto implement
* Simplest RAID storage subsystem design

**Disadvantages of RAID-1**

* Inefficient - 100% parity overhead is thehighest of all RAID types.
* Becomes very costly as number of disksincrease, it requires twice the desired diskspace
* The RAID function is done by systemsoftware, loading the CPU/Server anddegrading throughput at high activitylevels. Hardware RAID recommended
* May not support hot swap of failed diskwhen implemented in "software"

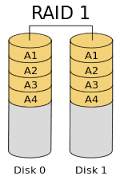


Figure 2. RAID-1

**RAID-3:**

RAID 3, which is rarely used in practice, consists of byte-level striping with a dedicated parity disk. One of the characteristics of RAID 3 is that it generally cannot service multiple requests simultaneously, which happens because any single block of data will, by definition, be spread across all members of the set and will reside in the same location.[clarification needed] Therefore, any I/O operation requires activity on every disk and usually requires synchronized spindles.

**Minimum number of drives required: 3**

**Recommended Applications:**

* Involves in large sequential data access such as Video Streaming

**Advantages of RAID-3:**

* It provides good bandwidth for the transfer of large volumes of data.

**Disadvantages of RAID-3:**

* This technology is fairly complex and too resource intensive to be done in software.
* Performance is slower for random, small I/O operations.

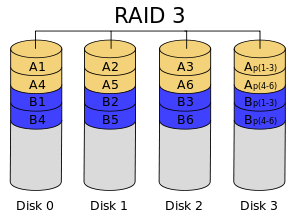


Diagram of a RAID 3 setup of six-byte blocks and two parity bytes, shown are two blocks of data in different colors.

**RAID-5:**

RAID 5 consists of 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.

In comparison to RAID 4, RAID 5's distributed parity evens out the stress of a dedicated parity disk among all RAID members. Additionally, read performance is increased since all RAID members participate in serving of the read requests.

**Minimum number of drives required: 3**

**Recommended Applications:**

* File and Application servers
* Database servers
* WWW, E-mail, and News servers
* Intranet servers
* Mostversatile RAID level

**Advantages of RAID-5:**

* High efficiency - highest read datatransaction rates, Medium Write datatransaction rates
* Good aggregate transfer rate
* Cost effective - only 1 extra disk isrequired
* Fault tolerant
* Low ratio of ECC (Parity) disks to datadisks means high efficiency
* The best choice is in multi-userenvironments which are not writeperformance sensitive.

**Disadvantages of RAID-5:**

* Disk failure has a medium impact onthroughput
* Most complex controller design
* Difficult to rebuild in the event of a diskfailure (as compared to RAID-1)
* Individual block data transfer rate sameas single disk

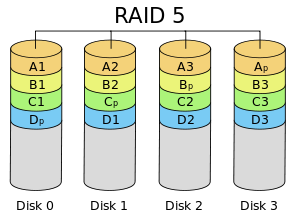
****

Diagram of a RAID 5 setup with distributed parity with each color representing the group of blocks in the respective parity block (a stripe). This diagram shows left asymmetric algorithm

**RAID 0+1:**

RAID-01 is technically a combination of RAID-1and RAID-0, includes both mirroring andstriping, but without parity. RAID-10 is a stripeacross a number of mirrored drives, and isimplemented as a striped array whosesegments are RAID-1 arrays. RAID-10 has thesame fault tolerance as RAID-1, as well as thesame overhead for fault-tolerance as mirroringalone. Advantages: Very high I/O rates areachieved by striping RAID-1 segments excellentsolution for sites that would normally use RAID-1 Great for Oracle and other databases whichneed high performance and fault tolerance.

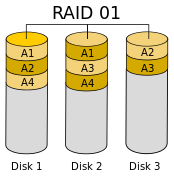
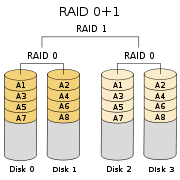
**Minimum number of drives required: 4**

**Advantages of RAID 0+1:**

* Fault tolerant
* Very High I/O rates

**Disadvantages of RAID 0+1**

* Very expensive - Expensive to maintainAs with Raid-1 total capacity is equal tohalf of the total capacity of all disk in thearray.
* High overhead
* Very limited scalability



A nested RAID 01 configuration A hybrid RAID 01 configuration

**RAID-10 A STRIPE OF MIRRORS:**

RAID-10 is not RAID 0+1. RAID-10 uses RAID-1mirroring and RAID-0 striping, and bothsecurity and sequential performance. RAID-10is a striped RAID-0 array whose segments aremirrored RAID-1. It is similar in performance toRAID 0+1, but with better fault tolerance andrebuild performance. It has the same faulttolerance as RAID-1 with the same overhead forfault tolerance as mirroring alone. Typically fourplus hard drives are used, because RAID-10creates two pairs of mirrored arrays andcombines these arrays to form one RAID-0array.

**Minimum number of drives required: 4**

**Recommended Applications:**

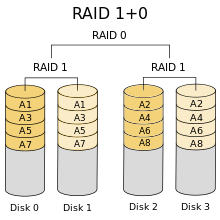
* Database server requiring highperformance and fault tolerance

**Advantages of RAID-10:**

* High fault tolerance
* High I/O rates achieved by striping RAID1 segments
* Faster rebuild performance than RAID0+1
* Under certain circumstances, RAID-10array can sustain multiple simultaneousdrive failures excellent solution for sites that would haveotherwise gone with RAID-1 but need someadditional performance boost.

**Disadvantages of RAID-10:**

* Very expensive, High overhead
* All drives must move in parallel to propertrack lowering sustained performance
* Very limited scalability at a very highinherent cost



A typical RAID 10 configuration

**Main difference between RAID 10 vs RAID 01**

Performance on both RAID 10 and RAID 01 will be the same.

The storage capacity on these will be the same.

The main difference is the fault tolerance level. On most implememntations of RAID controllers, RAID 01 fault tolerance is less. On RAID 01, since we have only two groups of RAID 0, if two drives (one in each group) fails, the entire RAID 01 will fail. In the above RAID 01 diagram, if Disk 1 and Disk 4 fails, both the groups will be down. So, the whole RAID 01 will fail.

RAID 10 fault tolerance is more. On RAID 10, since there are many groups (as the individual group is only two disks), even if three disks fails (one in each group), the RAID 10 is still functional. In the above RAID 10 example, even if Disk 1, Disk 3, Disk 5 fails, the RAID 10 will still be functional.

So, given a choice between RAID 10 and RAID 01, always choose RAID 10.

**RAID-6: Dual Parity Stripes**

RAID-6 is essentially an extension of RAID-5 which allows for additionalfault tolerance by using a second independent distributed parity scheme(dual parity).Data is striped on a block level across a set of drives, as in RAID-5.A second set of parity is calculated and written across all the drives.

**Minimum number of drives required: 4**

**RAID-6 Advantages:**

* It provides for an extremely high data fault tolerance and cansustain multiple simultaneous drive failures.
* It protects against multiple bad block failures while non-degraded.
* It protects against a single bad block failure while operating in adegraded mode.

**Raid-6 Disadvantages:**

* More complex controller design.
* Controller overhead to compute parity addresses is extremely high.
* Write performance can be brought on par with RAID-5 by using a customASIC for computing Reed-Solomon parity.
* Requires N+2 drives to implement because of dual parity scheme.

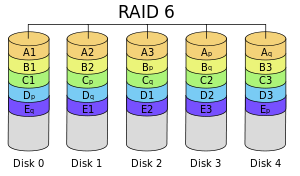


Diagram of a RAID 6 setup, which is identical to RAID 5 other than the addition of a second parity block

**Conclusion:** Thus, we studied various RAID levels with its advantages and disadvantages successfully.

**AIM: IMPLEMENTATION OF RAID LEVELS**

**Pre-requisites:**

* Any Unix based OS installed primarily or virtually.
* Package – “mdadm” should be installed on this system. This is ‘Multiple Device Admin Tool’. Acts as a device driver for implementing software RAID.

**Step 1.**

Create partitions on your system using following command.

**$ fdisk /dev/sda**

* n (for new)
* p (for primary partition)
* 1 (No. of partition)
* +200M (size)

Using this create 3 to 4 partitions. You will now have sda1, sda2, sda3 and so on.

* w (to write the changes)

Now run:

**$ sudo partprobe /dev/sda**

(to make the changes effective)

Now change the type of your partitions.

**$ fdisk /dev/sda**

* t (for type)
* fd (hex code for RAID type)

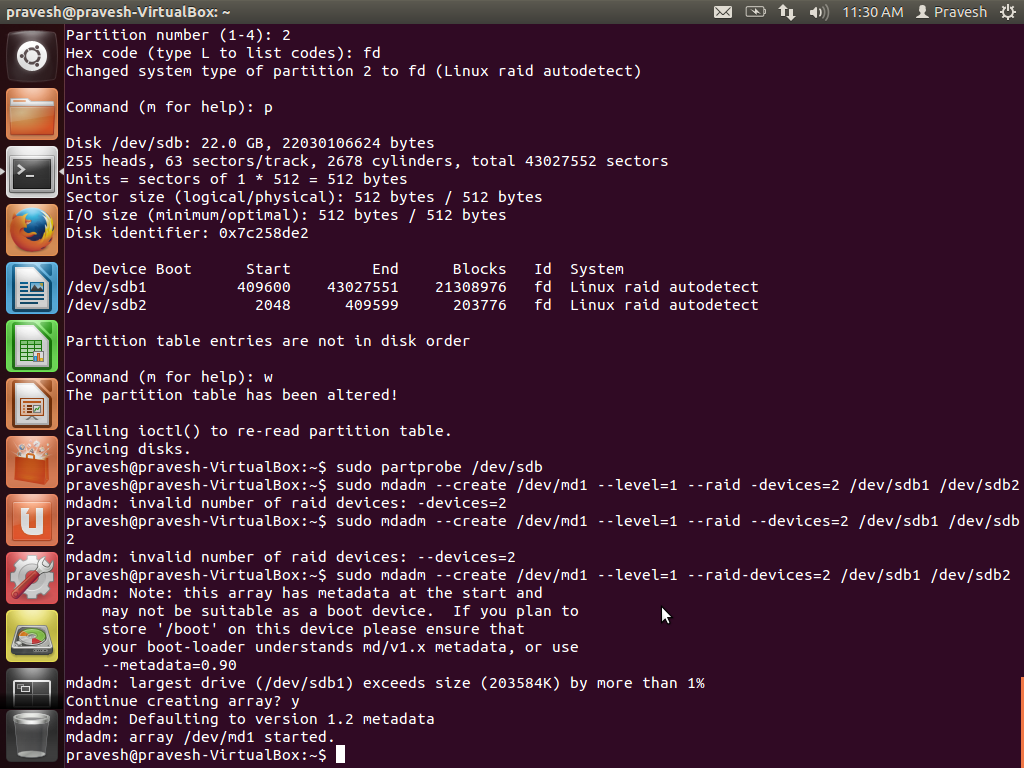
Then write the changes.

**Step 2.**

Now we create the RAID levels using following command.

**$ sudo mdadm –create /dev/md1 –level=1 –raid-devices=2 /dev/sda1/sda2**

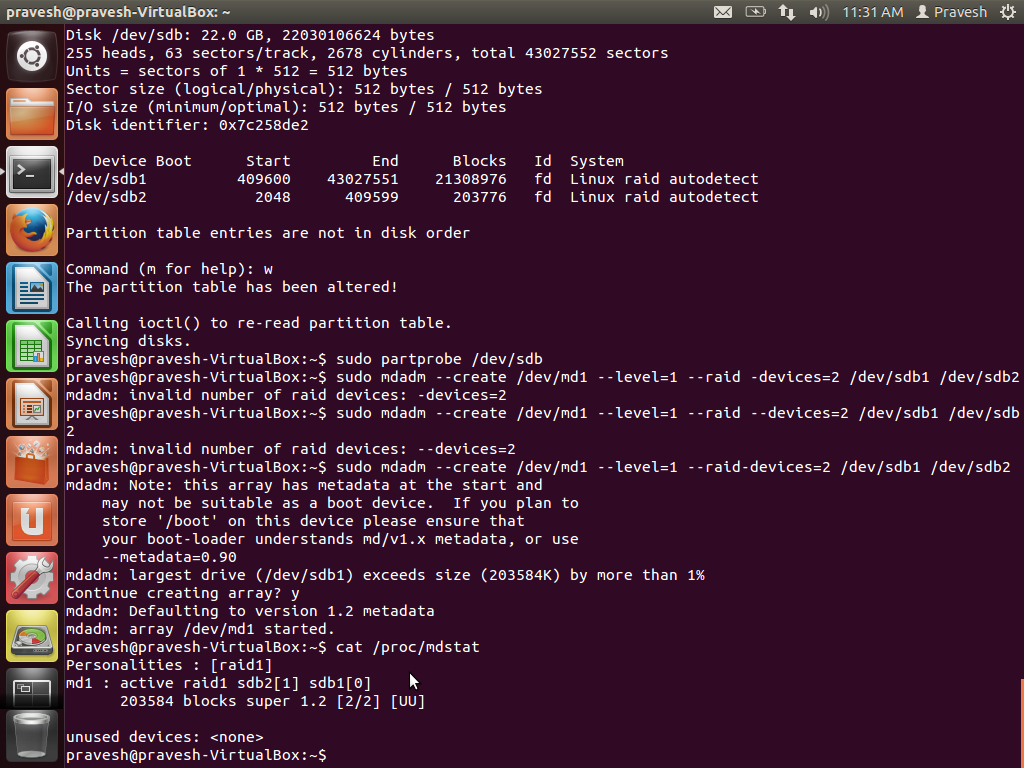
Here we are going to implement RAID 1.



Now to display the status of the RAID levels:

**$ cat proc /proc/mdstat**

This shows how many RAID partitions are active.

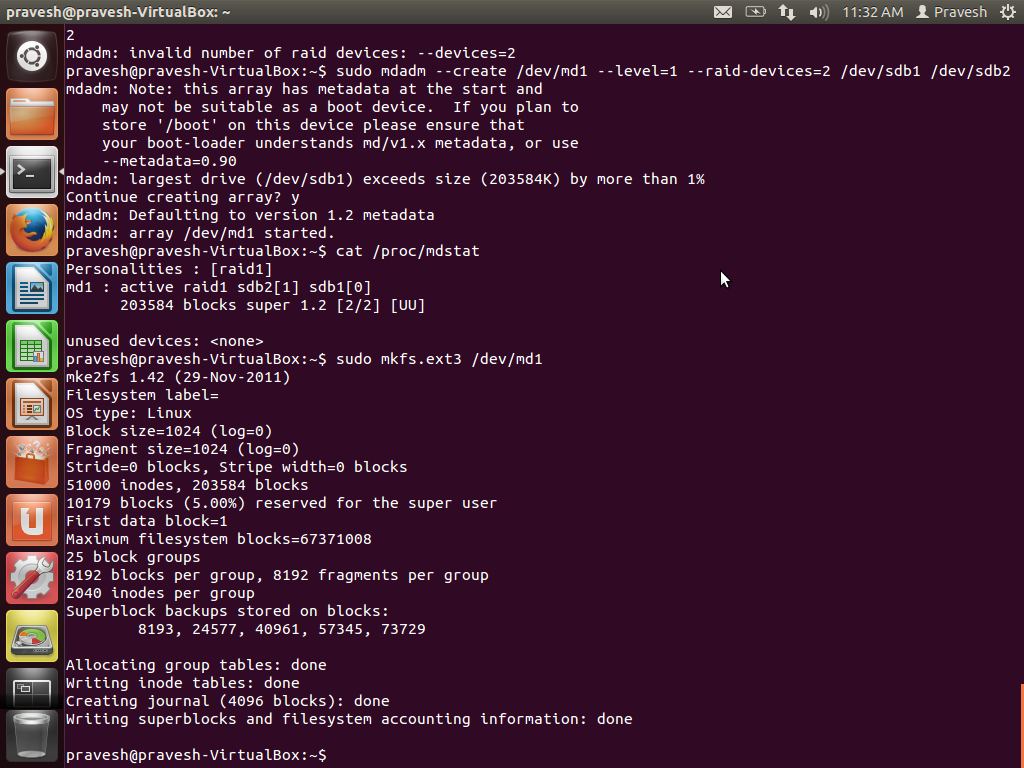


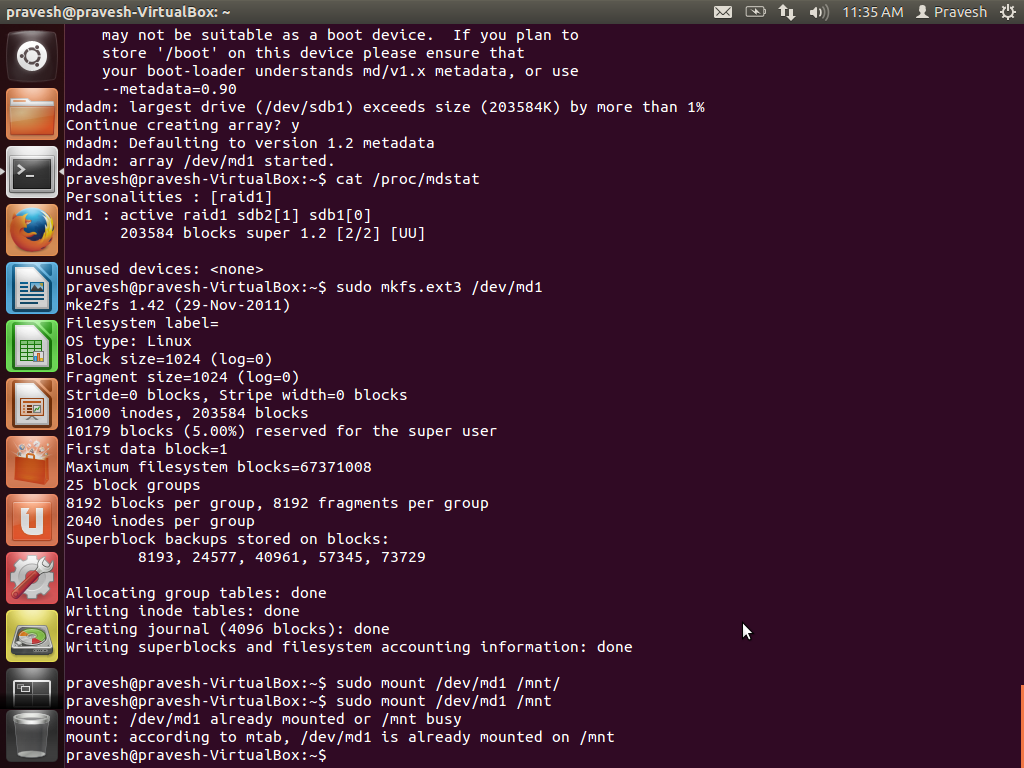
**Step 3.**

To mount the volumes so that we can actually use them.

**$ sudo mkfs.ext3 /dev/md1**

**$ sudo mount /dev/md1/mnt/**

****



**Conclusion:** Thus we have implemented RAID levels.

**Conclusion:** Thus we have implemented RAID levels