Project report on …..

STARGET

A storage assistant for you ….!!!

-Documented by :
Amaani S Karimi
Balaji Pawade
Khushal Chandak

CONTENT **Page no.**

1) OVERVIEW

This project is providing a utility specific to system administrators using shell scripting which will provide them to do storage configuration in easy way without affecting other configurations and removing the overhead of remembering different commands .
Under this we will be covering the topics such as :
providing quotas
Taking backups
Configuration of raid
Configuration of logical volumes
Monitoring the size and data percentage of exhaustion of logical volume
Metadata recovery

2) PROJECT PLAN

We will be managing certain task such as:

2.1) BASIC MONITORING

As soon as we login to client this utility will give detail information of ip, hostname, uptime, kernel, process, disk storage, file system information, detail lv information, lv and vg name with size and consumption, network information - dns, server, mac and gateway

2.2) QUOTAS

Disk space can be restricted by implementing disk quotas which alert a system administrator before a user consumes too much disk space or a partition becomes full.

Disk quotas can be configured for individual users as well as user groups. This makes it possible to manage the space allocated for user-specific files (such as email) separately from the space allocated to the projects a user works on (assuming the projects are given their own groups).

In addition, quotas can be set not just to control the number of disk blocks consumed but to control the number of inodes (data structures that contain information about files in UNIX file systems).Because inodes are used to contain file-related information, this allows control over the number of files that can be created.

-soft limit
-hardlimit
-grace period

2.3) BACKUPS

The purpose of most backups is to create a copy of data so that a particular file or application may be restored after data is lost, corrupted, deleted or a disaster strikes.
Backing up data is sometimes confused with archiving data, although these operations are different. A backup is a secondary copy of data used for data protection.

2.3.1) Full backup :- As the name implies, this type of backup makes a copy. The primary advantage to performing a full backup during every operation is that a complete copy of all data is available with a single set of media.
However, the disadvantages are that it takes longer to perform a full backup than other types (sometimes by a factor of 10 or more), and it requires more storage space.

2.3.2) Incremental backup :- An incremental backup operation will result in copying only the data that has changed since the last backup operation of any type.
The benefit of an incremental backup is that they copy a smaller amount of data than a full. Thus, these operations will complete faster, and require less media to store the backup.

The disadvantage is that during a complete restore, the latest full backup and all subsequent incremental backups must be restored, which can take significantly longer than restoring a full backup.

2.3.3) Differential backup :- A differential backup is a cumulative backup of *all* changes made since the last *full* backup, i.e., the *differences* since the last full backup. The advantage to this is the quicker recovery time, requiring only a full backup and the last differential backup to restore the entire data repository. The disadvantage is that for each day elapsed since the last full backup, more data needs to be backed up, especially if a significant proportion of the data has changed, thus increasing backup time as compared to the incremental backup method.A differential backup refers to a backup made to include the differences since the last full backup.

2.4) RAID

RAID means redundant array of independent disk .The basic idea is to combine multiple small, disk drives into an array to accomplish performance or redundancy goals not attainable with one large and expensive drive. This array of drives appears to the computer as a single logical storage unit or drive.

RAID allows information to be spread across several disks. RAID uses techniques such as disk striping(RAID Level 0)
disk mirroring(RAID Level 1)
disk striping with parity(RAID Level 5)
To achieve redundancy, lower latency, increased bandwidth, and maximized ability to recover from hard disk crashes

System Administrators and others who manage large amounts of data would benefit from using RAID technology. Primary reasons to deploy RAID include:
Enhances speed
Increases storage capacity using a single virtual disk

Minimizes data loss from disk failure

2.4.1) Stripping

RAID 0, also known as disk striping, is a technique that breaks up a file and spreads the data across all the disk drives in a RAID group.
The benefit of RAID 0 is that it improves performance. Because striping spreads data across more physical drives, multiple disks can access the contents of a file, allowing writes and reads to be completed more quickly. A drawback to RAID 0 is that it does not have parity. If a drive should fail, there is no redundancy and all data would be lost.

2.4.2) Mirroring

Disk mirroring, also known as RAID 1, is the replication of data to two or more disks. Disk mirroring is a good choice for applications that require high performance and high availability, such as transactional applications, email and operating systems.
Because both disks are operational, data can be read from them simultaneously, which makes read operations quite fast. The RAID array will operate if one disk is operational. Write operations, however, are slower because every write operation is done twice.
Disk mirroring provides instantaneous failover for data required by mission-critical applications. If primary arrays are damaged, traffic is switched to secondary or mirrored backup arrays.

2.4.3) Parity

parity (from the Latin *paritas*, meaning equal or equivalent) is a technique that checks whether data has been lost or written over when it is moved from one place in storage to another or when it is transmitted between computers.
Because data transmission is not an entirely error-free process, data is not always received in the same way as it was transmitted. A parity bit adds checksums into data that enable the target device to determine whether the data was received correctly.
An additional binary digit, the *parity bit,* is added to a group of bits that are moved together.

2.5) LOGICAL VOLUME

The **Volume Group** is the highest level abstraction used within the LVM. It gathers together a collection of Logical **Volumes** and Physical **Volumes** into one administrative unit.
What is logical volume manager and lvm2 ?
LVM is a tool for logical volume management which includes allocating disks, striping, mirroring and resizing logical volumes.

With LVM, a hard drive or set of hard drives is allocated to one or more *physical volumes*. LVM physical volumes can be placed on other block devices which might span two or more disks.

The physical volumes are combined into *logical volumes*, with the exception of the /boot/ partition. The /boot/ partition cannot be on a logical volume group because the boot loader cannot read it. If the root (/) partition is on a logical volume, create a separate /boot/ partition which is not a part of a volume group.

Since a physical volume cannot span over multiple drives, to span over more than one drive, create one or more physical volumes per drive.

The volume groups can be divided into *logical volumes*, which are assigned mount points, such as /home and / and file system types, such as ext2 or ext3. When "partitions" reach their full capacity, free space from the volume group can be added to the logical volume to increase the size of the partition. When a new hard drive is added to the system, it can be added to the volume group, and partitions that are logical volumes can be increased in size. On the other hand, if a system is partitioned with the ext3 file system, the hard drive is divided into partitions of defined sizes. If a partition becomes full, it is not easy to expand the size of the partition. Even if the partition is moved to another hard drive, the original hard drive space has to be reallocated as a different partition or not used.

LVM2-

LVM version 2, or LVM2, is the default for Red Hat Enterprise Linux 5, which uses the device mapper driver contained in the 2.6 kernel. LVM2 can be upgraded from versions of Red Hat Enterprise Linux running the 2.4 kernel.

steps required to configure LVM include:

Creating *physical volumes* from the hard drives.
Creating *volume groups* from the physical volumes.
Creating *logical volumes* from the volume groups and assign the logical volumes mount points

2.5.1) Thin provisioning

Thin provisioning involves using virtualization technology to give the appearance of having more physical resources than are actually available. If a system always has enough resource to simultaneously support all of the virtualized resources, then it is not thin provisioned. The term thin provisioning is applied to disk layer in this article, but could refer to an allocation scheme for any resource. For example, real memory in a computer is typically thin-provisioned to running tasks with some form of address translation technology doing the virtualization. Each task acts as if it has real memory allocated. The sum of the allocated virtual memory assigned to tasks typically exceeds the total of real memory.
Thin provisioning offers more efficiency where the amount of resource used is much smaller than allocated, so that the benefit of providing only the resource needed exceeds the cost of the virtualization technology used.Thin provisioning is a mechanism that applies to large-scale centralized computer disk-storage systems, SANs, and storage virtualization systems. Thin provisioning allows space to be easily allocated to servers, on a just-enough and just-in-time basis. Thin provisioning is called "sparse volumes" in some contexts.
Some of the important parameters and attributes which are kind of
Thin Pool - You can then create devices that can be bound to the **thin pool** for later allocation when an application actually writes to the logical volume

Passdown - pass down Process discards in the thin pool (as with no pass down), and pass the discards down the the underlying device. This is the default mode.

Fstrim - Removing files in a file system on top of a thin LV does not generally add free space back to the thin pool. Manually running the fstrim command can return space back to the thin pool that had been used by removed files. fstrim uses discards and will not work if the thin pool LV has discards mode set to ignore.

AUTOEXTEND -
dmeventd should be configured to extend thin pool LVs before all data space is used. Warnings are emitted through syslog when the use of a thin pool reaches 80%, 85%, 90% and 95%. The point at which dmeventd extends thin pool LVs, and the amount are controlled with two configuration settings found in /etc/lvm/lvm.conf
thin_pool_autoextend_threshold
is a percentage full value that defines when the thin pool LV should be extended. Setting this to 100 disables automatic extension. The minimum value is 50.
thin_pool_autoextend_percent
defines how much extra data space should be added to the thin pool LV from the VG, in percent of its current size.

Disabling
There are multiple ways that extension of thin pools could be prevented:

- If the dmeventd daemon is not running, no monitoring or automatic extension will occur.Even when dmeventd is running, all monitoring can be disabled with the lvm.conf monitoring setting.
- To activate or create a thin pool LV without interacting with dmeventd, the --ignoremonitoring option can be used. With this option, the command will not ask dmeventd to monitor the thin pool LV.
- Setting thin_pool_autoextend_threshould to 100 disables automatic extension of thin pool LVs, even if they are being monitored by dmeventd.

Data space exhaustion

When properly managed, thin pool data space should be extended before it is all used . If thin pool data space is already exhausted, it can still be extended

2.5.2) Snapshots

Thin snapshot volumes allow many virtual devices to be stored on the same data volume. This simplifies administration and allows for the sharing of data between snapshot volumes.
As for all LVM snapshot volumes, as well as all thin volumes, thin snapshot volumes are not supported across the nodes in a cluster. The snapshot volume must be exclusively activated on only one cluster node.

Thin snapshot volumes provide the following benefits:
A thin snapshot volume can reduce disk usage when there are multiple snapshots of the same origin volume.

If there are multiple snapshots of the same origin, then a write to the origin will cause one COW operation to preserve the data. Increasing the number of snapshots of the origin should yield no major slowdown.
Thin snapshot volumes can be used as a logical volume origin for another snapshot. This allows for an arbitrary depth of recursive snapshots (snapshots of snapshots of snapshots...).

A snapshot of a thin logical volume also creates a thin logical volume. This consumes no data space until a COW operation is required, or until the snapshot itself is written.

A thin snapshot volume does not need to be activated with its origin, so a user may have only the origin active while there are many inactive snapshot volumes of the origin.

When you delete the origin of a thinly-provisioned snapshot volume, each snapshot of that origin volume becomes an independent thinly-provisioned volume. This means that instead of merging a snapshot with its origin volume, you may choose to delete the origin volume and then create a new thinly-provisioned snapshot using that independent volume as the origin volume for the new snapshot.
Although there are many advantages to using thin snapshot volumes, there are some use cases for which the older LVM snapshot volume feature may be more appropriate to your needs:
You cannot change the chunk size of a thin pool. If the thin pool has a large chunk size (for example, 1MB) and you require a short-living snapshot for which a chunk size that large is not efficient, you may elect to use the older snapshot feature.

You cannot limit the size of a thin snapshot volume; the snapshot will use all of the space in the thin pool, if necessary. This may not be appropriate for your needs.

2.6) Metadata Recovery

The configuration details of a volume group are referred to as the metadata. By default, an identical copy of the metadata is maintained in every metadata area in every physical volume within the volume group. LVM volume group metadata is stored as ASCII.

By default, the pvcreate command places the physical volume label in the 2nd 512-byte sector. This label can optionally be placed in any of the first four sectors, since the LVM tools that scan for a physical volume label check the first 4 sectors. The physical volume label begins with the string LABELONE.

LVM Metadata

1) Configuration details of a volume group.

2) Stored as ASCII format directly on disk.

3) Identical copy of the metadata is maintained in every metadata area in every physical volume within the volume group.

4) The volume group metadata contains:
Information about how and when it was created.
Information about the volume group.

5) Latest metadata file is present in /etc/lvm/backup/vg-name.

Failure scenarios

LVM command will throw errors Couldn't find device with uuid ...Couldn't find all physical volumes for volume group ,Checksum error Couldn't read volume group metadata. Volume group metadata is inconsistent

LVM devices are not present but can be seen in mount Output

Probable causes

- Overwritten metadata
- Corrupted metadata
- accidentally deleted either VG or some LVs Restore metadata
- Partition table alteration

Recovery -

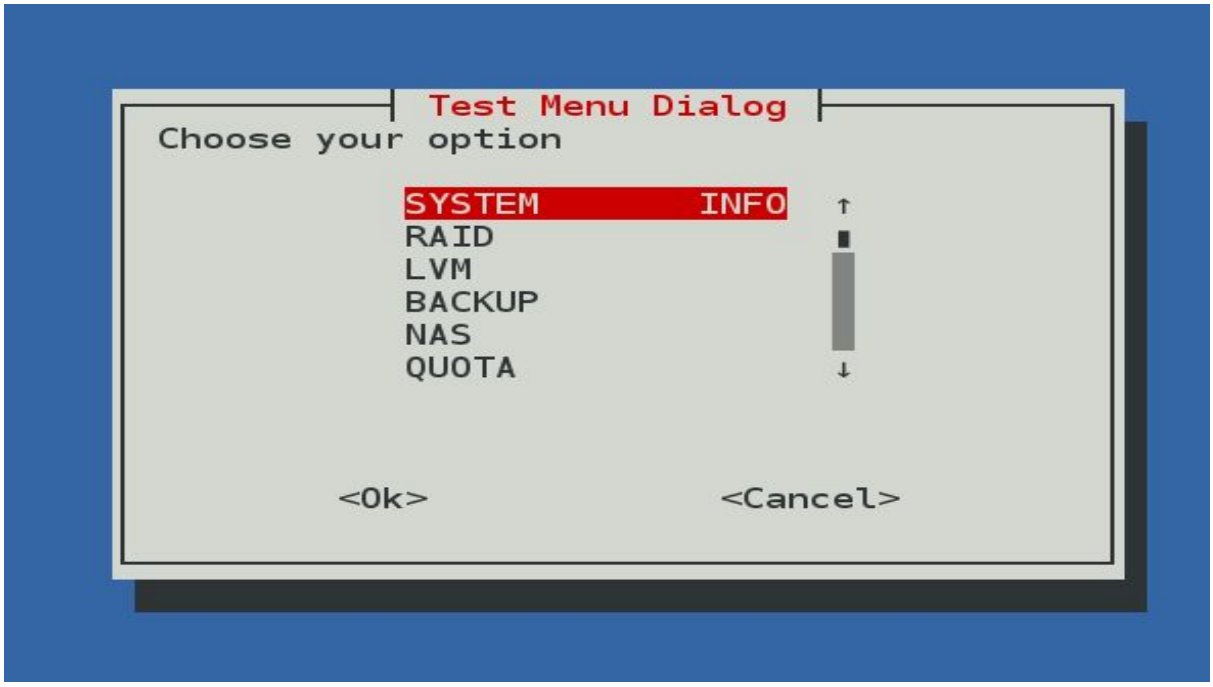
If the volume group metadata area of a physical volume is accidentally overwritten or otherwise destroyed, or that the system was unable to find a physical volume with a particular UUID. You may be able to recover the data the physical volume by writing a new metadata area on the physical volume specifying the same UUID as the lost metadata.

3) PROJECT REQUIREMENTS

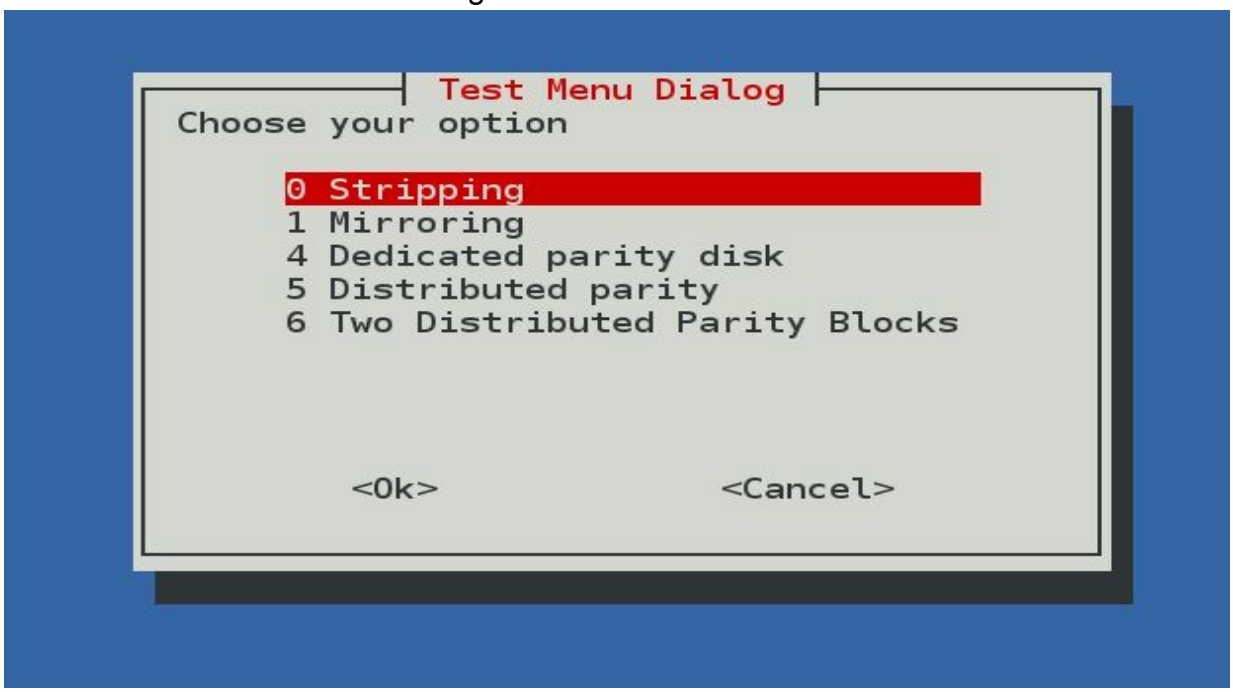
1 packages
2 rhel 7 operating system

4) Design and Screenshot

starget is a tui based utility and it will provide the interface as shown in above picture.

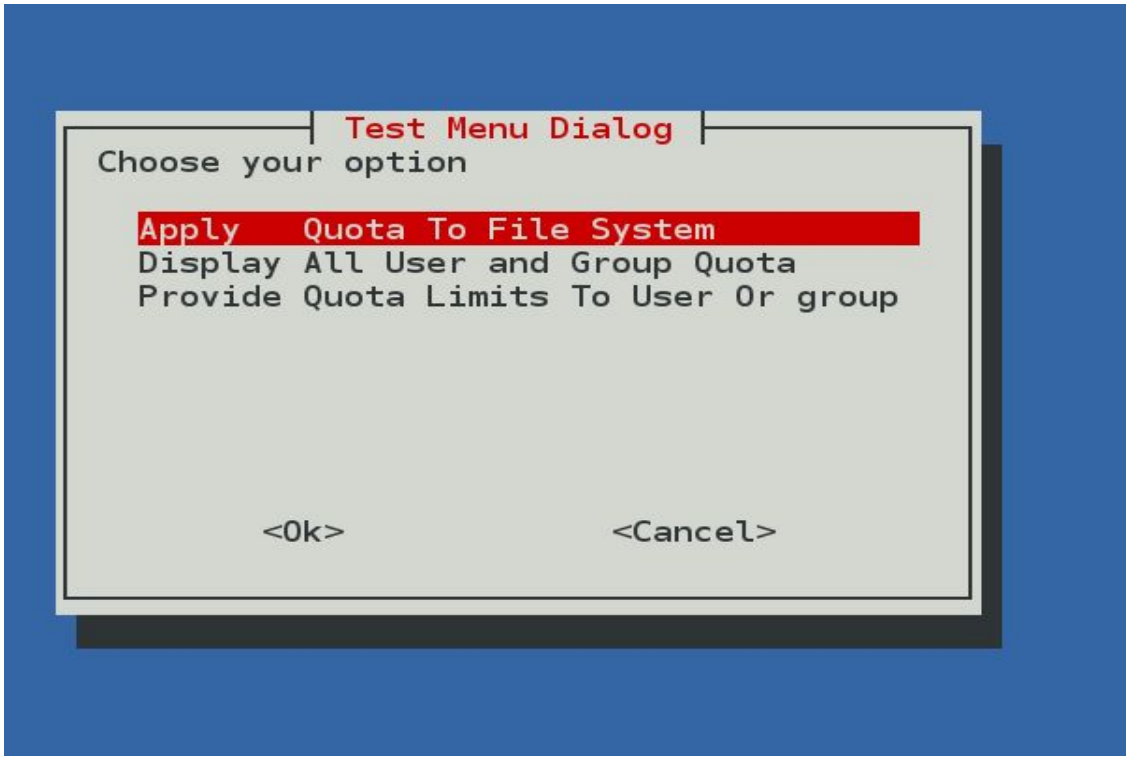


After selecting RAID as an output following options will be prompted to the user for raid levels-
Here starget will display all raid compatible disk on your system by filtering through all disk which is free for further configuration.

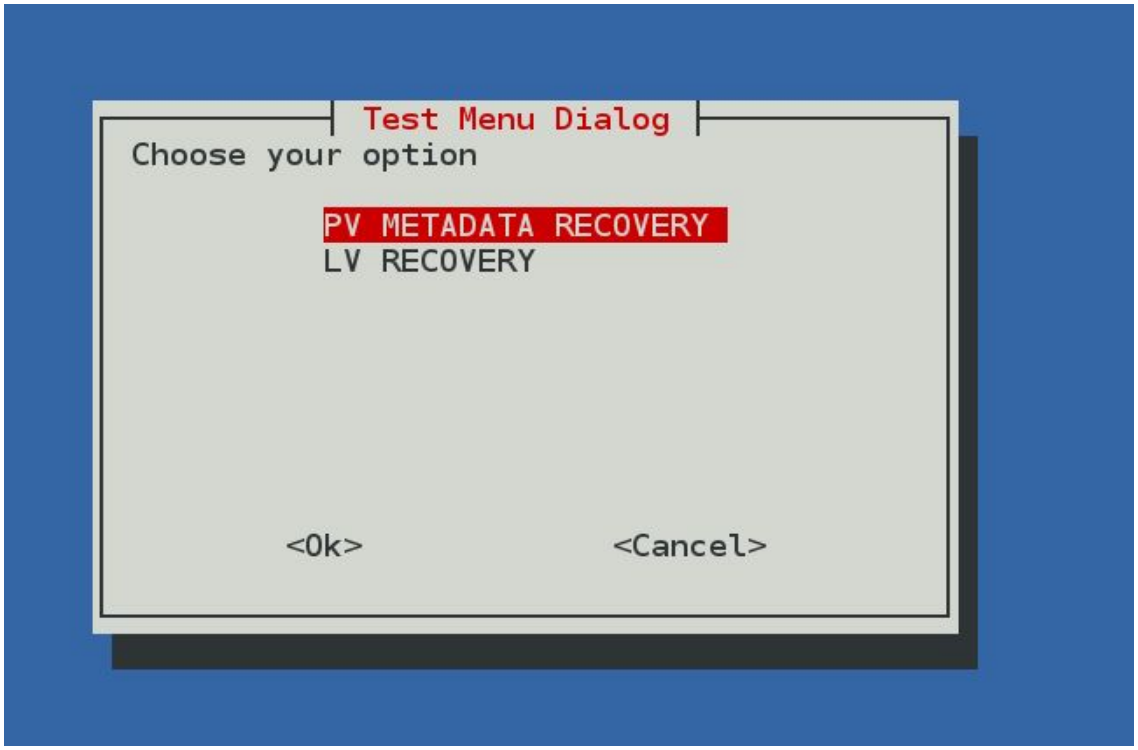


Backup option provides this interface with three options - complete , differential and incremental.

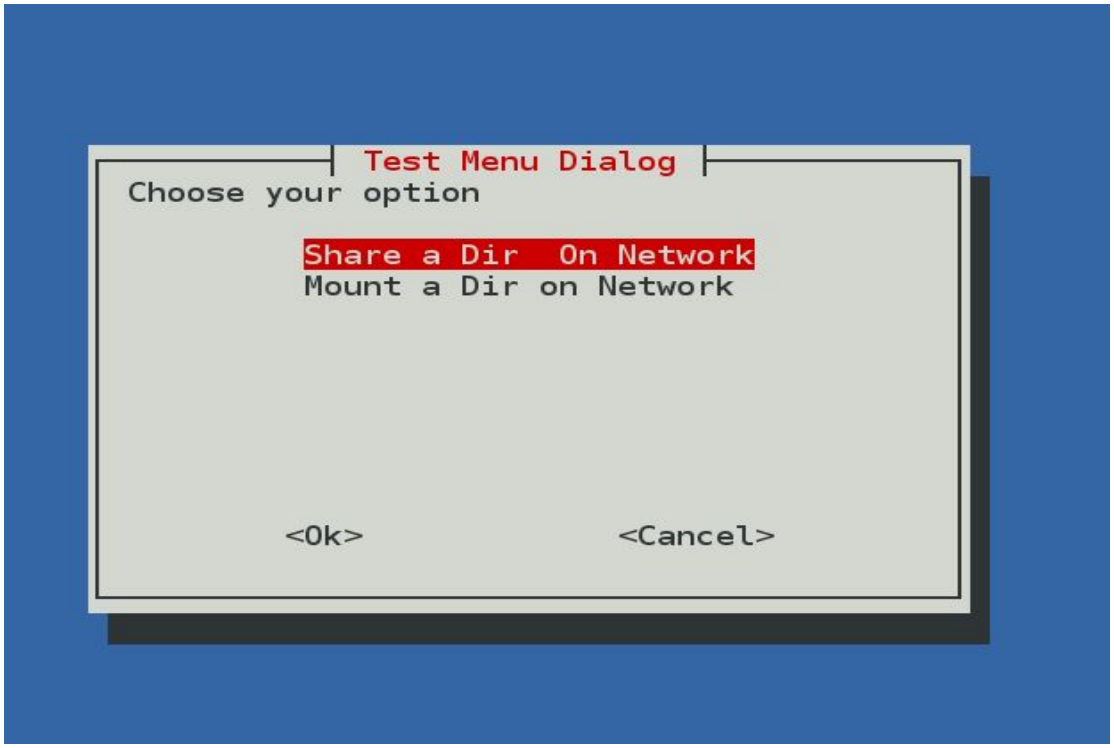
Quotas allocation provide the feature of applying quotas with soft and hard limit also provides certain grace interval also admin can display and see all applied quotas till that instant .



Since metadata is very crucial part for data recovery of all important data ,our product will bring back the data by recovering the corrupted metadata or recovering your logical volume back to its working state.



Below , is the screenshot of how an interface of NAS will look like , it provides all read , write and sync permission access to the file you choose also, you can mount a directory on network as we usually do in nfs.



5) script-full (Link to see implementation)

https://drive.google.com/drive/folders/0B4Ni8q0QuAhkbF9TUnlha0tuM2s?usp=sharing

6) Limitation

- Limited to single system environment
- Does not cover all meta data recovery scenarios
- Some validation is still remaining
- Snapshot is ready but not yet implemented
- Quotas on xfs is working fine but will include when 100% sure on that .

7) Scope

- Xfs quotas allocation .
- Snapshot with more precision .
- Using lvm filters and more detailed attribute .
- Cover other leftover meta data recovery scenarios.
- Restoration of the backup .