

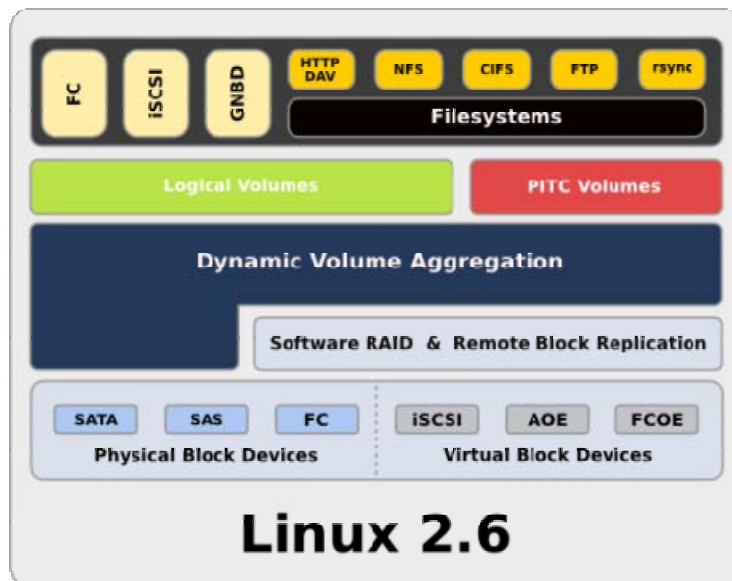
Architecture Overview

ezNAS consolidates several open source technologies on the Linux 2.6 kernel base to deliver a comprehensive storage management solution that meets the needs of enterprise applications, administrators and users.

- **Operating System** - Linux 2.6 is at the core of the storage-optimized ezNAS Operating System Environment. Linux 2.6 is a stable, high-performance, enterprise-class Operating System.
- **Block Devices** - any disk or raw volume is considered a block device in ezNAS. These can be physical block devices attached directly to the ezNAS host appliance, or remote block devices - designated virtual block devices - located remotely of the ezNAS host appliance. Supported physical block devices are SATA, SAS, SCSI and FC disks. Aggregated physical block devices such as RAID volumes also fall within this designation. In addition, ezNAS can work with remote block devices to present to the upper layers of the ezNAS storage architecture. Remote block devices accessed via iSCSI, AOE or FCOE protocols are defined as virtual block devices. The ezNAS dynamic volume aggregation and software RAID & replication layers make no distinction between physical block devices and virtual block devices.



EDR Host	Type	Description	Size	Label Type	Partitions
...	SCSI	Virtual, VMware Virtual S	1.00 TB	none	1 LUN
...	SCSI	Virtual, VMware Virtual S	1.00 TB MB	RAID	1 LUN
...	SCSI	Virtual, VMware Virtual S	1.00 TB MB	RAID	1 LUN
...	SCSI	Virtual, VMware Virtual S	1.00 TB MB	RAID	1 LUN
...	SCSI	Virtual, VMware Virtual S	1.00 TB MB	RAID	1 LUN
...	SCSI	Virtual, VMware Virtual S	1.00 TB MB	RAID	1 LUN
...	SCSI	Virtual, VMware Virtual S	1.00 TB MB	RAID	1 LUN



- **Software RAID & Remote Block Replication** - ezNAS supports block device aggregation using RAID protocols via a software RAID management layer. Building a disaster-recovery solution is also very easy with ezNAS support for synchronous and asynchronous block-level replication. Block-level replication is

done over standard TCP/IP protocol, allowing for data transfer to occur over local area and wide area networks.

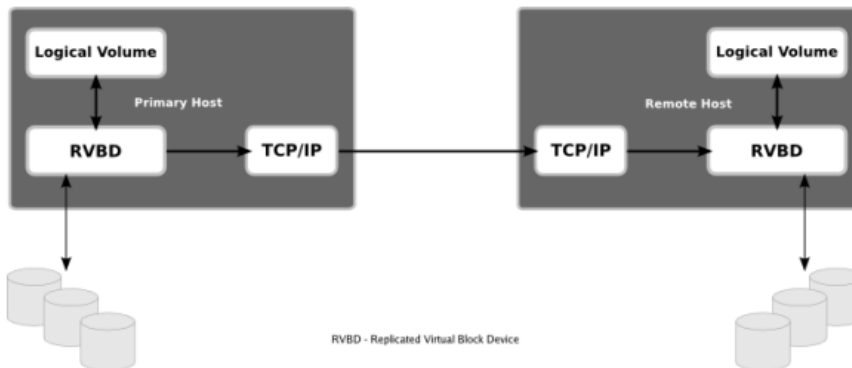
- Software RAID** - ezNAS software RAID allows the administrator to aggregate physical and virtual block devices to increase usable capacity, enhance block I/O performance and provide extended reliability and availability. Multiple software RAID arrays are supported, and any block device - physical or virtual - can participate as a RAID member. ezNAS even supports RAID on RAID. Multiple physical RAID volumes can be members of a software RAID volume. Administrators can choose from several RAID algorithms to fit specific application deployment scenarios. ezNAS supports RAID 0, RAID 1, RAID 5, RAID 6 and RAID 10.



- Remote Block Replication** - one can never underestimate the value of a sound disaster recovery plan. ezNAS meets this challenge with support for remote block replication. Synchronous and asynchronous replication can



be achieved over a standard TCP/IP network connection. Synchronous replication is perfect when the secondary host where blocks are copied to can be reached via a high-bandwidth connection; gigabit+. For a remote host on a slower connection, say in a different building or even country, asynchronous replication ensures that I/O performance is not compromised at the expense of data security. A replicated block device in ezNAS is presented to the upper block device management layer as a virtual block device. The replication happens transparently of the I/O operations that network clients make to the ezNAS server.



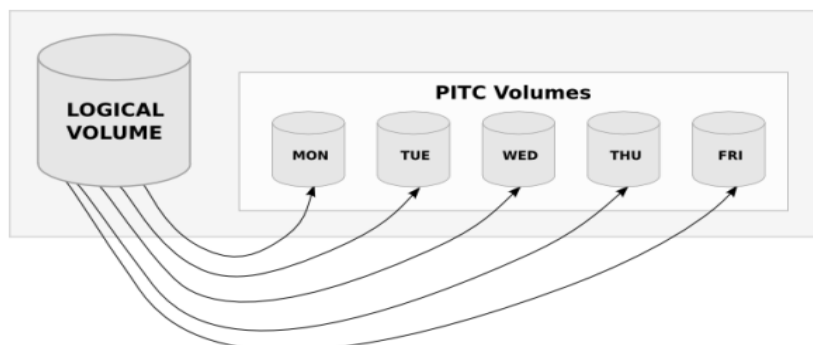
Any data written to the replicated block device via any of the supported storage export protocols eventually finds its way to the remote host. Depending on whether replication is done synchronously or asynchronously, I/O performance on the client side will vary. Irrespective of the replication mode, all data reads from the client will be serviced from the local host, while writes are not complete until acknowledged by the replication layer.

- **Dynamic Volume Aggregation** - before storage is presented to the export layers, a pool must be created of all the raw block storage; virtual block devices, RAID volumes and replicated volumes are concatenated into a single storage pool. This storage pool can then be further sliced to be allocated as network block exports or filesystem shares. Dynamic volume aggregation also confers the ability to create point in time copies of sliced volumes, which can be used as part of a comprehensive backup strategy. With point in time copies, or snapshots, I/O can continue on source volumes while backups of PITC volumes progresses.

Volume Group Management							
Member PVs of VG "transient"		Size	Allocated	Free	Members	Add physical storage	Delete VG
Device	Size	2.91 GB	0 bytes	2.91 GB	View member PVs	All PVs are used	Delete
/dev/sdc1	1017.98 MB	2.97 GB	0 bytes	2.97 GB	View member PVs	All PVs are used	Delete

Think of DVA as an abstraction layer that groups disparate block devices and virtualizes them to be used as logical volumes or partitions. ezNAS supports multiple DVA instances which allows the administrator to group block devices according to performance and/or availability metrics. For instance, an appliance that has both high capacity SATA disks and high performance SAS disks can have two DVA instances; a SATA grouping for backup and archiving, and a SAS grouping for database and application I/O.

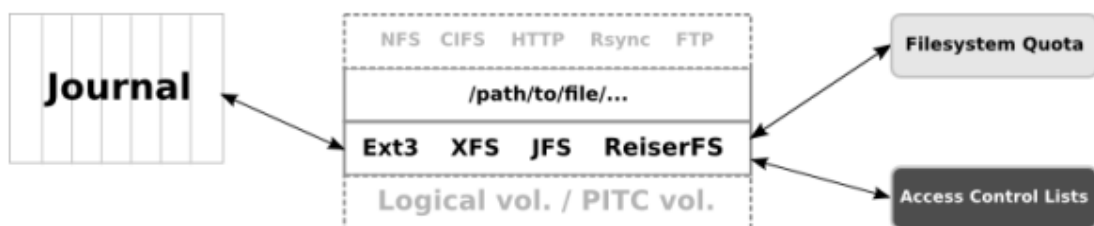
- **Logical Volumes & PITC Volumes** - ezNAS allows the administrator to apportion slices or partitions of a DVA instance which can then be presented to



clients over any of the supported storage networking protocols. These slices are referred to as logical volumes - and they in effect provide a second level of indirection between the exported data and the physical or virtual block device which that data is stored on. This allows transparent manipulation of the underlying physical block devices without affecting the structure of the exported volume.

A Point in Time Copy volume is a frozen image of a logical volume. A logical volume can have one or more PITC volumes, representing the state of the logical volume at the time the PITC volume is created. With a PITC volume, I/O on its source logical volume can continue uninterrupted, while a consistent backup of the PITC volume takes place. A PITC volume can also be exported, and accessed by network clients, exactly like a regular logical volume.

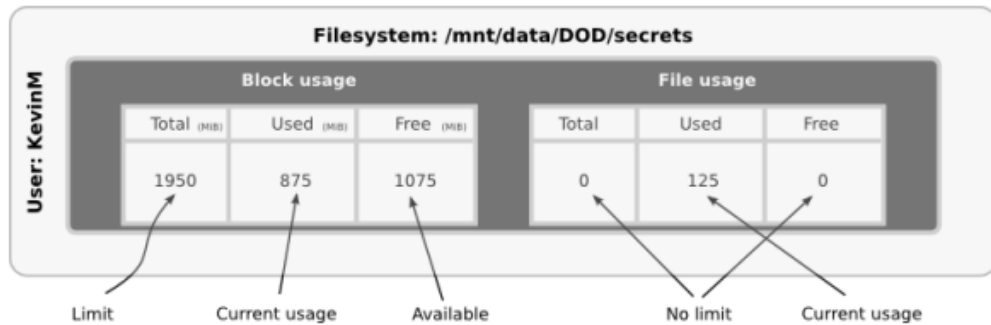
- **Filesystems** - ezNAS supports multiple on-disk filesystems. Administrators can create a logical volume and apply a filesystem that meets the needs of specific application I/O profiles or performance parameters. The two major filesystems supported by ezNAS are XFS and ext3. Both these filesystems are journaled filesystems. A journaled filesystem enhances data security and reduces the need for regular filesystem checks. Any data written to disk is first logged to a journal before being committed. This mechanism ensures data consistency in the event of a system crash or power outage. Other filesystems supported by ezNAS are ReiserFS v3 and JFS. ezNAS will also soon support GFS, the journaled clustering filesystem. ext3 filesystems can be as large as 8TB and XFS can scale to 10's of terabytes on 64-bit architecture.



Two fundamental user and group management mechanisms factor in the filesystem level. *Quota management* and *Access Control Lists*.

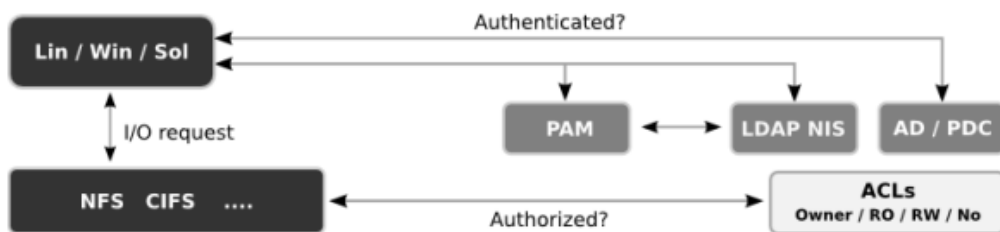
- **Quota** - Administrators can restrict usage of disk capacity on a per-filesystem basis. Restrictions can be made both at the group level and user level. This granularity allows the administrator to assign blanket parameters to a group and subsequently restrict the disk resources of users within that group. Usage quota and file quota may be assigned via the management interface. Usage quota restricts the amount of disk capacity

taken up by the management object - be it a group or a user. Restrictions can also be placed on the actual number of filesystem objects, files and



directories, that a user or group may consume. The administrator may use either or both restriction mechanisms for controlling the available storage resource.

- Access Control Lists** - Network clients accessing storage via the file-based network storage protocols such as CIFS and NFS may do so in guest mode or authenticated mode. With guest mode access, no authentication or authorization for the client is required before being granted access to storage on the ezNAS appliance. In authenticated mode, clients need to first be authenticated via one of the supported authentication mechanisms. Once authenticated, clients will then need to be authorized for access to data via an access control mechanism. This access control mechanism, access control lists (ACL), is a function of the filesystem on which the data is stored. Each filesystem object - a file or directory - has a set of ACL metadata objects assigned to it that determines whether a user or a group has access to that filesystem object.



A unique design feature in ezNAS in the implementation of ACL management is the ability to configure ACLs at the group level only. This has implications for exporting storage over file-level storage networking protocols such as CIFS. Generally, this feature ensures that administrators are not overly taxed with creating and assigning access control configurations for individual users, but rather only to groups containing



those users. Another advantage of this ACL allocation method is that performance degradation on the filesystem does not occur as a result of the ACL filtering algorithm. Each filesystem has a physical limit on the number of ACL metadata objects per filesystem object. For ext3, that limit is 32. This restriction could be a determining factor in choosing what filesystem to use for specific storage deployments. It is likely however that future versions of ezNAS will converge on a single filesystem, most likely XFS. Any new deployments should take this aspect into consideration.

- **Service Layer** - at the top of the stack lies the service layer. This is where storage is exported via block-level or file-level protocols. Clients access storage via CIFS, NFS, HTTP/DAV, FTP or rsync at the file-level; and Fibre Channel, iSCSI or GNBD at the block level. File-level protocols are layered over a filesystem whereas block-level protocols feed directly off logical volumes. Both block-level and file-level protocols support exporting PITC volumes.

ezNAS is a family of products that consolidates several open source technologies on the Linux 2.6 kernel base. These ezNAS solutions extend and modify the capabilities of the popular open source program Openfiler for specific integration with Xiotech's Intelligent Storage Element based Emprise 5000. Therefore, not all capabilities and features are enabled on all ezNAS products.