

2. Опишете ZFS

The **Z file system (ZFS)**, originally developed by Sun Microsystems, is designed to use a pooled storage method in that space is only used as it is needed for data storage. It is also designed for maximum data integrity, supporting data snapshots, multiple copies, data checksums and is easy to administrate. It uses a software data replication model, known as RAID-Z. RAID-Z provides redundancy similar to hardware RAID, but is designed to prevent data write corruption and to overcome some of the limitations of hardware RAID.

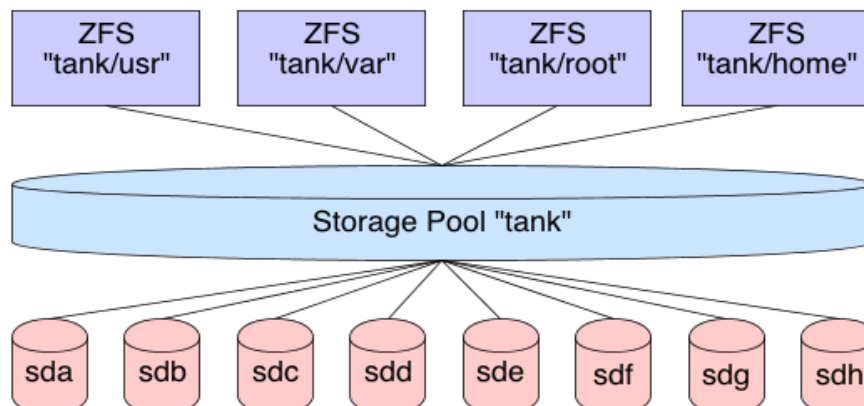
Storage pool

ZFS uses the concept of storage pools to manage physical storage. File systems were constructed on top of a single physical device (see the figure below).



To address multiple devices and provide for data redundancy, the concept of a volume manager was introduced to provide a representation of a single device so that file systems would not need to be modified to take advantage of multiple devices. This design added another layer of complexity and ultimately prevented certain file system advances because the file system had no control over the physical placement of data on the virtualized volumes.

ZFS eliminates volume management. Instead of creating virtualized volumes, ZFS aggregates devices into a storage pool. The storage pool describes the physical characteristics of the storage (device layout, data redundancy, and so on) and acts as an arbitrary data store from which file systems can be created. File systems are no longer constrained to individual devices, allowing them to share disk space with all file systems in the pool. You no longer need to predetermine the size of a file system, as file systems grow automatically within the disk space allocated to the storage pool. When new storage is added, all file systems within the pool can immediately use the additional disk space without additional work. In many ways, the storage pool works similarly to a virtual memory system.



For example, if we have 80 GB hard disk and we use 1 GB for swap, 10 GB for / and 69 GB for /home. After few months the storage for / is not enough anymore. We put another storage device in our PC and what are we doing then? If we have another file system, we should do some additional work, loose time etc. With ZFS it is not needed. The / will take the needed space from /home and we shouldn't do any additional changes, the file system do it for us.

Checksums

The user data is verified using a user-selectable checksum algorithm. All checksum verification and data recovery are performed at the file system layer, and are transparent to applications.

Snapshots

A snapshot is a read-only copy of a file system or volume. Snapshots can be created quickly and easily. Initially, snapshots consume no additional disk space within the pool.

Scalability

The ZFS is 128 bit file system and supports 256 quadrillion zettabytes.

RAID-Z

RAID-Z avoids the RAID 5 write hole using copy-on-write: rather than overwriting data, it writes to a new location and then automatically overwrites the pointer to the old data. It avoids the need for read-modify-write operations for small writes by only ever performing full-stripe writes. Small blocks are mirrored instead of parity protected, which is possible because the file system is aware of the underlying storage structure and can allocate extra space if necessary.