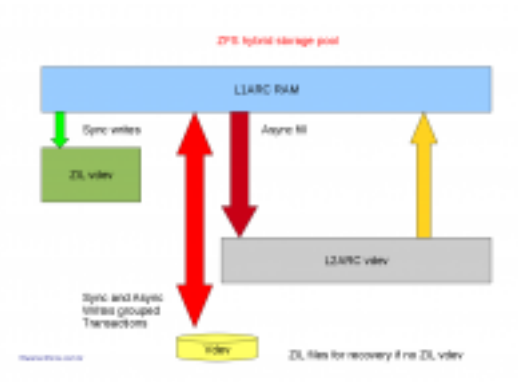
ZFS parte 3: quick presentation for sysadmins

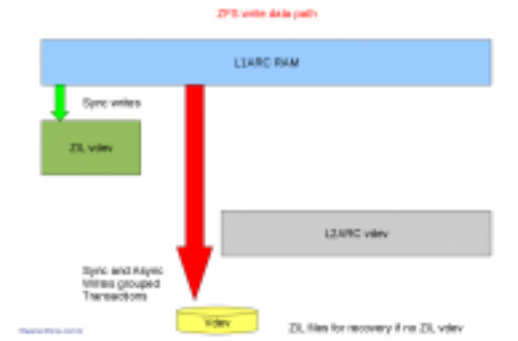
Author: André Felipe Machado<andremachado@techforce.com.br>

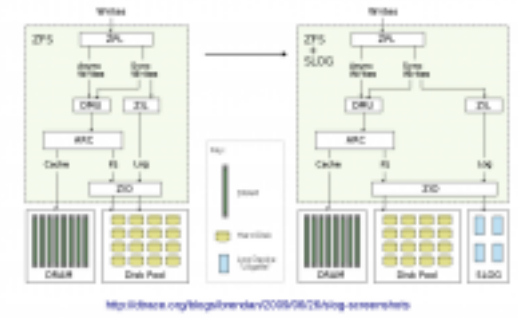
This material was used for a ZFS features and some internals quick presentation for system administrators. You could talk for seconds to hours of details at each topic and graphic element. Send suggestions or corrections for improvement. There is a PDF link ready for use in presentations.

Some ZFS key features

- Self healing: write and read checksum against bit rot.
- Copy on write transactional.
- Online scrub (“fsck”) and resilvering (“raid”). Online resizing.
- Compression: LZJB, gzip1-9, ZLE .
- Deduplication.
- Unlimited snapshots and clones.
- Hybrid storage pool: auto tiered.
- Variable block size: 512 B to 128 KB, dynamic when compression enabled.
- Serializes and aggregate writes in transaction groups, TXG.
- “Unlimited” file size (16 EB).
- 128 bit scalability, without inode pre-allocation.
- 2^48 files per directory.
- Thin provisioning datasets.
- File and block protocols: CIFS, NFS, iSCSI. FC using ComStar stmfadm.
- Includes block device and volume management: RAID-0, RAID-Z, Z2, Z3, mirror, ZVOL,VDEV, pool, hot spares.
- Remote replication.
- Multiple filesystems and mountpoints into one pool.
- Number of file copies beyond RAIDZ or mirror redundancies.
- User and group quotas dataset filesystems.
- Per dataset filesystems space reservations in one pool (“thick provisioning”).
- RAID-Z* does not have RAID write hole.
- JBOD using SAS/SATA expanders instead of detrimental RAID controllers.
- ARC with ZIL VDEV reduce fragmentation.







http://dtrace.org/blogs/brendan/2009/08/28/slog-screenshots

In a future article we will write about ZFS IOPS and latency.

zfs_hsp_write_read_datapath