Center for
Materials
Technologies

Storage Technologies

Skolkovo
Institute of Science
and Technology

Skoltech

Storage technologies: brief summary

Technology	Maturity	Price	Consistency	Supported host OS
Hardware RAID	Mature	High	None	In theory, any. In practice, Linux (kernel support) and Windows (vendor-specific drivers) only.
mdadm+LVM	Mature	Zero	None	Linux only
ReFS+StorageSpaces	Young	Moderate	High	Windows only
Btrfs	Young	Zero	High	Linux only
ZFS	Mature	Zero	High	Solaris[1], FreeBSD[2], MacOS[3], Linux[4]

Technology	Access to physical drives	Simplicity of monitoring
Hardware RAID	Host OS does not have access to physical drives.	Hard to monitor. Vendor-specific tools required.
mdadm+LVM	Required. Physical drives must be accessible from the host OS (motherboard ports or HBA controller).	Open source tools available (e.g. smartmontools).
ReFS+StorageSpaces		
btrfs		
ZFS		

Notes:
[1] – Origin.
[2] – First port. Very mature.
[3] – Currently not supported on bare metal.
[4] – Thanks to LLNL and OpenZFS team.

RAID levels

Level	Redundancy	Pros	Cons
RAID 0 (stripe)	None	Size = N * DriveSize (where N is number of drives). Maximum possible performance.	Zero fault tolerance, should only be used for “zero-price” data.
RAID 1 (mirror)	It depends	If N > 2, then redundancy is maximum possible. High read performance.	Size = DriveSize Size always equals to the size of single drive.
RAID 10 (stripe of mirrors)	Moderate	Size = N * DriveSize / 2 High performance.	In theory, redundancy is lost after the failure of one drive (in practice, when N is large, after such a failure only “Russian roulette” starts).
RAID 5	Moderate	Size = (N – 1) * DriveSize	Minimal redundancy (but still fault tolerant). Should not be used when N ~ 7 or more.
RAID 6	Always high	Size = (N – 2) * DriveSize	In case of N ~ 12 -- none. Should not be used when N ~ 20 or more.

Terminology.
“Hot spare” is the failover mechanism to provide extra reliability.
“Hot spare drive” is the drive which is automatically switched into operation in case of active drive failure.

Note: some technologies (ZFS at least) allow to create exotic configurations, e.g. RAID 50 (stripe of RAID 5), RAID 60 or even RAID7.

Choices for DELL R750xs with PERC 755

Hardware state: seven front panel disks connected to PERC 755 (instead of PERC 355 which was requested by Skoltech).

PERC 755 **can only operate in RAID** (not HBA) **mode**.
As a consequence, **only one technology can be used**: **hardware RAID** (see first line on slide 2).
The most advanced technology – ZFS – **is not available** (ZFS **as file system can still be used**, providing high consistency, snapshots, clones, high speed compression, etc.)

Three variants of configuration can be proposed.

1. **Minimal** redundancy: RAID 5 of all 7 disks. Available size **22 TB**.
2. **Moderate** redundancy: RAID 6 of all 7 disks. Available size **18 TB**.
3. **Maximal** redundancy: RAID 6 of 6 disks plus hot spare. Available size **14 TB**.

Note: in case of ZFS over RAID (possible, though suboptimal configuration) and fast lz4 compression, the effective available space can be approximately doubled.

Thx!

Skoltech