

Lecture 14

RAID management

New context

- Cloud computing: storage as a service
- Extremely large capacity
- Reliability
- Fast access to data
- Fault tolerance
- Scalability
- Easy to manage
- Transparent administration
- Secure access to data

Reliability via redundancy

- Data striping: splitting the bits of each byte across multiple disks – bit-level striping. This can be generalised to include a number of disks that is either a multiple of or divides 8.
- RAID level 0: striping at the level of blocks but without any redundancy.
- RAID level 1: disk mirroring.
- RAID level 2: uses error-correcting code (ECC). One disk store the 1st bit of all bytes, another disk the 2nd bit,... Two/three disks store the error-correcting bits.

- RAID level 3: bit-interleaved parity organisation, improves on level 2 by using only one disk for parity.
- RAID level 4: block-interleaved parity organisation, keeps a parity block on a separate disk. In one of the disks fails, the parity block can be used with the corresponding blocks from the other disks to restore the blocks on the failed disk.
- RAID level 5: block-interleaved distributed parity, spreads data and parity among all disks.

Implementation of RAID

- In the kernel, e.g. RAID 0, 1, or 0+1
- In the host-bus adapter hardware. It's not flexible.
- In the hardware of the storage array. The OS needs to implement the file system on each of the volumes.
- In the storage area network layer by disk virtualisation

Linux RAID support

- 2.6 kernel supports md: arrays can be built on top of entire disks or on partitions.
- mdadm is now the standard RAID management tool and should be found in any modern distribution.
Download from
<http://www.cse.unsw.edu.au/~neilb/source/mdadm/>
- mdadm has 7 major modes of operation. Normal operation just uses the 'Create', 'Assemble' and 'Monitor' commands.
- [/proc/mdstat](#): the array is running. you can create a file system, just like you would on any other device, mount it, include it in your /etc/fstab, and so on.

Questions

- Could RAID level 1 achieve better performance for read requests than RAID 0 (with non-redundant striping of data) ? If so, how ?
- Consider a RAID 5 architecture with five disks; the fifth stores the parity block. How many blocks are accessed in order to perform the following ?
 - A write of one data block;
 - A write of seven continuous blocks of data.