Oscar Petersson, Matteus Laurent

oscpe262, matla782

# TDDI41 Lab report

Högskoleingenjörsutbildning i datateknik, 180 hp

**Exercise 1: Very basic RAID theory**

**RAID** originally stood for *"Redundant Array of Independent Disks"*, but nowadays, it is commonly known as *"Array of Independent Disks"*, i.e. a collection of storage media working as a single logical unit.

**JBOD** (*"Just a Bunch Of Disks"*) is multiple drives exposed as individual devices. Sometimes, the conception of spanned disks is expressed as 'JBOD'.

**RAID-0** (*"Striping"*) distributes the content (near-) equally across all disks to increase throughput. Any disk failure results in a full array failure.

**RAID-1** (*"Mirroring"*) writes the data set identically to two drives, basically allowing a "backup" in case of one of the disks failuring.

**RAID-5** (*"Block-level striping with distributed parity"*) requires at least three disks, distributing data across two with parity on the third. This allows one disk to fail with the possibility to recover the array.

**RAID-6** (*"Block-level striping with double distributed parity"*) requires at least four disks, distributing data as RAID 5 but with parity on two disks.

**RAID 0+1** (*"Mirrored stripes"*) stripes data across two sets of disks which are in turn mirrored.

**RAID 1+0** (*"Stripe of mirrors"*) mirrors data and stripes the mirrored sets.

**Hot Spare** is a disk added to a failure redunant array, such as RAID-5, which is used only upon an array failure, at which point the array can start to rebuild instantly.

## Exercise 2: Volume management with LVM2

A physical disk is divided into one or more **Physical volumes** (`pv[command]`). Combining these makes **Volume groups** (`vg[command]`) which consists of any number of **Logical volumes** (`lv[command]`).

## Exercise 3: File Systems

The main differences between ext2 and ext3 is the addition of journaling (block and metadata) in ext3, and the possibility to grow a mounted file system on the latter.

Ext$n$ file systems are created with the command:
  `mkfs.ext`$n$ `<device>`

Ext file systems are resized with:   `resize2fs <device> <size>`

`fsck` is used to check and repair Linux file systems.

## Exercise 11: Some theory

**11-1 Why is RAID 1+0 generally considered more reliable than RAID 0+1?**

Lets assume we have six drives, (1-6). For RAID 0+1 we use disk 1-3 striped, and mirror the stripe to disks 4-6. For RAID 1+0 we mirror disk 1 to disk 2, 3 to 4, and 5 to 6, and stripe the

three sets.

Now, consider that we have two mirrored sets in RAID 0+1 and three in RAID 1+0. Should one disk in the 1-3 array fail in RAID 0+1 and one in the 4-6 during recovery, the array is doomed. However, for the RAID 1+0 array to fail, one disk in each set could fail and we'd still be able to recover the array, increasing the likelyhood of recovering the array before it is too late. So, even though the odds of two (contentwise) identical disks failing, the fault tolerance is increased.

## 11-2 As individual disk sizes grow, RAID-5 is no longer considered useful, and RAID-6 is also rapidly becoming obsolete. Why?

Mirrored setups, such as RAID 1+0, takes less time to recover. Parity schemes require all blocks of the drives in the array, whereas the the earlier only require those of the failed drive. This means that recovery of failed parity-based arrays are more likely to have a second (or third ...) failure during the recovery, rendering the array useless. All this assumes that read and write speeds does not follow the increase in disk sizes, but when did that happen last? ;)

## 11-3 What is the point of the file system journal in e.g. ext3?

One purpose is to reduce the amount of I/O operations by letting multiple operations be done in one go. Another purpose is to get a transaction behavoiur, is case of interruptions in the process.

## 11-4 Explain the purpose of using a volume manager, like LVM

We'd say there are two main purposes. The first is to allow disks to be added and replaced (hot swap implied) without downtime. The second is to allow snapshots on file systems which does not come with such functionality.

## 11-5 Optional: compute the probability of data loss over three years in a RAID-6 array consisting of 24 6TB SAS disks, with two hot spares (i.e. 22 disks in the RAID set). The RAID controller has 150MB/s throughput of which 80% can be used to rebuild. Assume MTBF of 145000 hours, latent sector error of 10E-13.

We'll get back to that if there is time ...