

Linux RAID, LVM

Ondřej Caletka



27. listopadu 2012



Uvedené dílo podléhá licenci Creative Commons Uveďte autora 3.0 Česko.

- 1 Teorie o RAID
- 2 Konfigurace softwarového RAIDu
- 3 Úvod do LVM
- 4 Práce s LVM

Odbočka o zarovnání oddílů

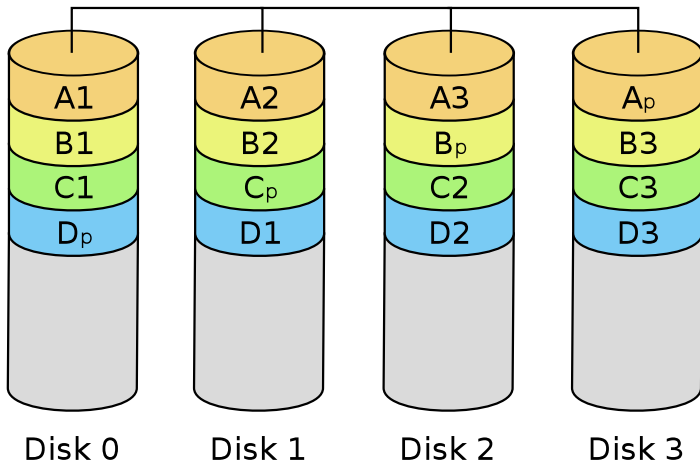
- Disky mají sektory délky 512 Bajtů.
- Modernější disky mají ale ve skutečnosti větší sektory.
- Požadavek na přenos sektoru 2, 3, 4 a 5 je mnohem pomalejší než požadavek na 0, 1, 2 a 3.
- Updatujte `fdisk` na takový, který zarovnává na sektory místo cylindrů.

Co je to RAID

- Redundant Array of Inexpensive Disks
Nadbytečná množina nenákladných disků
- Víc menších disků je levnější, než velký.
- Použitím n disků klesá n krát střední doba mezi poruchami.
- RAID umožňuje různými konfiguracemi nastavovat buď spíše velkokapacitní, nebo spíše zabezpečený systém.
- Skutečný RAID je HW podporovaný, pro OS vystupuje jako jeden velký bezporuchový disk.
- Linux podporuje SW RAID pomocí modulu MD, falšné RAID řadiče (fakeraid) pomocí dmraid.

- linear
 - pouze slepení dvou disků k sobě
 - LVM bude lepší volbou
- RAID-0
 - slepení disků s prokládaným ukládáním (striping)
- RAID-1
 - úplné záloha 1:1
 - disky vědí, když něco nevědí, takže nejsou potřeba tři
- RAID-5,6
 - částečná záloha s kapacitou $(n - 1)$, $(n - 2)$
 - distribuovaná parita

Princip RAID-5



Wikimedia commons, autor: Colin M.L. Burnett.

$$X_1 \oplus X_2 \oplus X_3 \oplus X_p = 0, \text{ kde } X=A, B, C, \dots$$

Vytvoření RAIDu

- Vytvoříme oddíly s *id = 0xfd* (Linux RAID autodetect)
- Sestavíme RAID pomocí
`mdadm --create /dev/md1 -l 5 -n 3 -e 0 /dev/sd[cde]1`
- Sledujeme inicializaci:
`cat /proc/mdstat`
`iostat -m`
- Alternativa: vybudujeme degradované pole – nahradíme jeden z disků slovem `missing`.
- Vytvořené zařízení `/dev/md1` můžeme dále dělit na oddíly, vytvářet souborové systémy, či použít jako LVM PV.
- Nastavení svazku je dobré uložit:
`mdadm --detail /dev/md1 --brief >> mdadm.conf`



- Deaktivace RAIDu (před vypnutím PC):

```
mdadm --stop /dev/md1
```

- Aktivace RAIDu (po zapnutí PC):

- `mdadm --assemble /dev/md1`

- provede automaticky kernel s autodetekcí

- pro RAIDy v konf. souboru: `--assemble --scan`

- Degradované spuštění: `--run`

- Details o RAIDu a o discích:

```
mdadm --detail /dev/md1
```

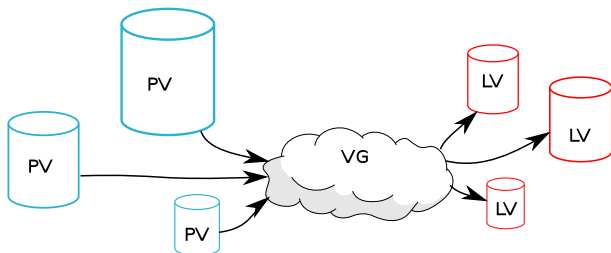
```
mdadm --examine /dev/sdc1
```


- Prohlášení disku vadným
 - `mdadm -f /dev/md1 /dev/sdc1`
 - Prove se automaticky, zmizí-li disk z aktivovaného RAIDu.
- Odebrání vadných disků
 - `mdadm -r /dev/md1 faulty`
- Přidání nového disku
 - `mdadm -a /dev/md1 /dev/sdf1`
 - Proběhne automaticky rebuild.
 - Proces rebuildu lze zrychlit pomocí bitové mapy.
- Odebrání RAID superbloku:
`mdadm --zero-superblock /dev/sdc1`

- Implementace RAIDu v linuxu umožňuje v omezené míře i rozšiřování existujících polí.
- Více od MD na <https://raid.wiki.kernel.org/>
- Přejít na RAID-1 krok za krokem:
http://www.linuxsoft.cz/article.php?id_article=723
- Velmi důležitý je trvalý dohled RAID pole.

Logical Volume Management

- Správa logických svazků
- Virtualizace blokových zařízení
- Implementováno pomocí Device Mapper
- Místo disků používáme logické svazky (LV).
- Tyto svazky alokujeme ve skupině svazků (VG).
- Poskytovatelem fyzického prostoru pro VG jsou fyzické svazky (PV).



- Výhody:
 - Snadná změna velikosti a počtu oddílů.
 - Nezávislost na připojovacím rozhraní – disky se indentifikují pomocí UUID.
 - Snapshoty – možnost kdykoli získat „hardcopy“ daného LV, zmrazenou v čase.
- Nevýhody:
 - Obvykle nezajišťuje redundanci dat.
 - Přidává další složitost – problémy se záchranou dat, instrukcí TRIM a root oddílem na LVM.

- Inicializace PV – oddíl s *id = 0x8e*:
`pvcreate /dev/sdc1`
- Inicializace VG – volba velikosti bloku (PE):
`vgcreate MyVG /dev/sd[cde]1`
- Vytvoření LV:
`lvcreate -n padesatmb -L 50m MyVG`
- Zjišťování informací:
`{pv,vg,lv}s`
`{pv,vg,lv}display`
`{pv,vg,lv}display -m` – vypíše mapování PE

- Změna velikosti LV:
`lvresize -L +10m /dev/MyVG/padesatmb`
Přepínač `-r` provede zároveň změnu velikosti filesystemu.
- Přesouvání dat mezi PV (za běhu):
`pvmove -vi 5 /dev/md1 /dev/sdc3`
- De/aktivace – potřeba provést při startu:
`vgchange -a [y/n]`

- Vytvoření snapshotu:
`lvcreate -s -L 30m -n snap1 /dev/MyVG/padesatmb`
- Velikost může být menší, než velikost původního oddílu, zapisují se jen změny.
- Snapshot je zapisovatelný.
- Snapshotu můžeme měnit velikost prostoru vyhrazeného pro změny.
- Snapshot zrušíme stejně jako LV:
`lvremove /dev/MyVG/snap1`

- Metadata jsou v textové podobě v `/etc/lvm/` a také na každém PV.
- LVM HOWTO je bohužel zastaralé.
- Dobrým zdrojem informací jsou články na serverech `root.cz` a `abclinuxu.cz`.

Děkuji za pozornost.

