Linux - pokročilá administrace

Roman Pavelka

Obsah

- Přehled a troubleshooting startu systému
- Práce s disky
- Zálohování systému
- Sledování událostí systému
- Sledování vytížení systému
- Disková pole RAID
- Logical Volume Manager (LVM)
- Nastavení ověřování přístupu uživatelů

Formát: 45 min / téma, 10 min pauza, 1 hodina na oběd

Prezentace dostupná zde: **romanpavelka.cz/kurzy**

Přehled a troubleshooting startu systému

- význam bootloaderu grub
- start jádra, prametry jádra
- koncept programu init
- služby spouštěné při startu
- Rescue režim
- Obnova zapomenutého hesla

Přehled a troubleshooting startu systému

Boot sequence:

1. CPU spustí UEFI (nebo BIOS)
2. Grafická karta inicalizuje obrazovku
3. Provede se RAM a PCI test
4. Načtou se disky a další bootovatelná zařízení
5. Spustí se bootloader
6. Bootloader spustí Operační systém

Přehled a troubleshooting startu systému

BIOS

MBR = Master boot record

512 B na prvním sektoru disku

Obsahuje:
zavaděč + partition tabulku

Struktura MBR

Adresa			Popis		Délka v bajtech
Hex	Oct	Dec			
0000	0000	0	Kód zavaděče		**440** (max 446)
01B8	0670	440	Volitelná signatura disku		4
01BC	0674	444	Obvykle nuly; 0x0000		2
01BE	0676	446	**Tabulka rozdělení disku (MPT)** (4 položky po 16 bajtech, IBM schéma oddílů)		**64**
01FE	0776	510	55_H	Signatura MBR; 0xAA55[1]	**2**
01FF	0777	511	AA_H		
Celková délka MBR: 446 + 64 + 2 =					**512**

Přehled a troubleshooting startu systému

Intel EFI, UEFI

Nepoužívá MBR
Co má spustit, je definováno na úrovni filesystému:

x86-64: \efi\boot\bootx64.efi

Virtual box používá defaultně BIOS metodu, ale lze nastavit EFI

Přehled a troubleshooting startu systému

Bootloader = program pro načtení operačního systému

GNU GRUB = GRand Unified Bootloader

Přehled a troubleshooting startu systému

Start jádra (Linux startup process)

1. Načtení a dekomprese jádra
2. Připojení disků
3. Načtení init RAM disku (initrd) s ovladači
4. Spuštění plánovače (scheduler)
5. Spuštění init procesu

Konfigurace jádra: /boot/config*

Přehled a troubleshooting startu systému

Koncept programu init

- První spuštěný proces, spouští ostatní procesy
- PID 1
- SystemV init, OpenRC, **systemd**, ...
- Runlevels, targets
  - SystemV: rc (run commands) skripty v /etc/rc*
  - systemd: system units v /etc/systemd/system
    - kompatibilitu s SystemV zajišťuje systemd-sysv-generator

Přehled a troubleshooting startu systému

systemd

- "system and service manager"
- přípona *d* znamená daemon
- Vedoucí vývoje Lennart Poettering (Red Hat)
- init systém, logování, správa uživatelů, zařízení a síťových připojení
- Konfigurační utilita systemctl

Přehled a troubleshooting startu systému

Služby spouštěné při startu (systemd):

```
systemctl list-unit-files | grep enabled
```

Přidání služby (zde webový server nginx)

```
systemctl enable nginx
```

Další příkazy:

```
status, start, restart, stop, disable, ...
```

Přehled a troubleshooting startu systému

Nouzový režim

rescue mode, recovery mode, single-user mode

Pro Debian je v grub Advanced menu

Alternativa: boot z USB flash disku + chroot, typicky pro:

- změnu zapomenutých hesel
- záchranu dat ze zcela nebootovatelného systému
- opravu poškozeného grubu

System Rescue (dříve System Rescue CD)

Práce s disky

- Kontrola konzistence
- Zálohování poškozeného média
- Vytváření kopií disků

Práce s disky

Výpis zařízení:

```
fdisk -l
```

Výpis připojených zařízení:

```
mount
```

Odpojení:

```
umount /dev/sdb1
```

Kontrola konzistence

```
fsck -l
```

Práce s disky

Fyzická kontrola

`badblocks`

(destruktivní i nedestruktivní testování čtení a zápisu)

Připojení disku

```
mount -t ext4 /dev/sdb1 /mnt/usb
```

Co se jak připojuje (mount)

`/etc/fstab`

Práce s disky

Zálohování poškozeného média

Pod úrovní filesystému:

```
dd if=/dev/defectiveDisk of=/target/disk/block-by-block.imh bs=1G conv=notrunc,noerror
```

Na úrovni filesystému:

```
rsync -av --ignore-errors /defective/disk /safe/disk
```

Práce s disky

Vytváření kopií disků

```
sudo dd if=/dev/sdb of=/dev/sdc status=progress
```

Zkopíruje vše, včetně prázdného místa!

Altenrativou je `dump`

Zálohování systému

- Práce s tar
- Zálohování po síti
- rdiff-backup

Zálohování systému

Práce s tar

Zabalení: `tar cf archive.tar folderA fileB`

Rozbalení: `tar xf archive.tar`

Zabalení s kompresí: `tar zcf archive.tar folderA fileB`

Zálohování systému

Zálohování po síti

Šifrovaný přenos:

```
scp -r SOURCE DEST
```

Optimalizovaný šifrovaný přenos (kopírují se jen rozdíly):

```
rsync -av --delete /Directory1/ /Directory2/
```

Zálohování systému

rdiff-backup

Záloha:

```
rdiff-backup <source directory> <backup directory>
```

Obnova:

rdiff-backup -r now `<backup directory>` `<source directory>`

- Komprese, statistiky, zachovává práva
- Zálohování změn

Zálohování systému

Pravidelné spouštění:

`cron`

`crontab -l`

`/etc/crontab`

`/etc/cron.*`

`/etc/cron.d/*`

Obrázek: crontab.guru

Sledování událostí systému

- možnosti sledování
- konfigurace syslog
- prohlížení logů
- nastavení automatické rotace log souborů

Sledování událostí systému

Možnosti sledování

`/var/log`

`dmesg, /var/log/messages`

systemd: `journalctl`

Sledování událostí systému

Konfigurace logování

```
/etc/systemd/journald.conf
```

```
systemctl reload systemd-journald
```

```
systemctl force-reload systemd-journald
```

Sledování událostí systému

Prohlížení logů (systemd)

```
journalctl

journalctl -b

journalctl -u nginx

journalctl --since 09:00 --until "1 hour ago"
```

Sledování událostí systému

Nastavení automatické rotace log souborů

Systemd:

`SystemMaxUse, SystemKeepFree, Storage=persistent/volatile`

Externí utilita logrotate:

`/etc/logrotate.conf`

\+ cron

less, cat, grep, ...

Pozor na časové zóny!

Sledování vytížení systému

- možnosti sledování vytížení základních částí systému
- procesor, paměť, disky, síť
- příkazy pro sledování
- interpretace výsledků
- sledování a ladění změnou systémových parametrů a konfigurace

Sledování vytížení systému

Možnosti sledování vytížení základních částí systému

`top`

`htop`

`/proc/`

`uname`

Sledování vytížení systému

Procesor:

`uptime, top, htop`

Paměť:

free -h

Disky:

`df -h`

Síť:

bmon

Grafické karty NVIDIA:

nvidia-smi

Sledování vytížení systému

Zabbix

Sledování vytížení systému

Interpretace výsledků

- RAM, swapping
- **HDD**
- File descriptors
- IO vs. CPU
- Paralelizace na CPU a GPU

Sledování vytížení systému

Sledování a ladění změnou systémových parametrů a konfigurace

```
sudo apt install tuned tuned-utils tuned-utils-systemtap
```

```
sudo tuned-adm profile throughput-performance
```

desktop, powersave, virtual-host, virtual-guest

noatime: filesystem performance

Disková pole RAID

- možnosti Linuxu pro zvýšení dostupnosti dat (High-Availability)
- vlastnosti diskových polí RAID
- HW a SW pole
- konfigurace RAID úrovní v Linuxu (mirror, stripe set apod.)
- sledování stavu
- obnova po havárii disku

Disková pole RAID

RAID = Redundant Array of Independent Disks

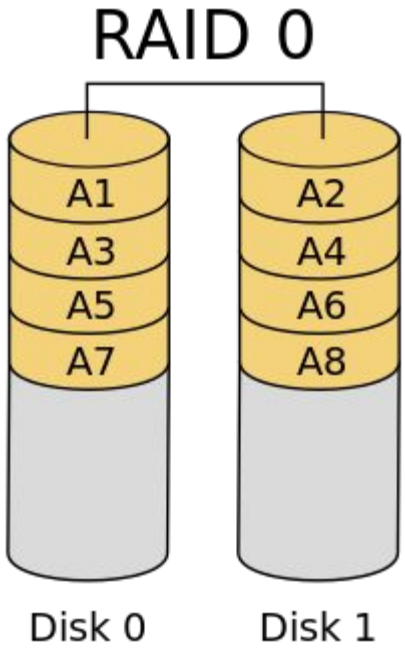
Motivace: spolehlivost a rychlost

Disková pole RAID

RAID 0: rychlost, horší spolehlivost

"stripping"

Při ztrátě jediného disku přijdeme o všechna data!

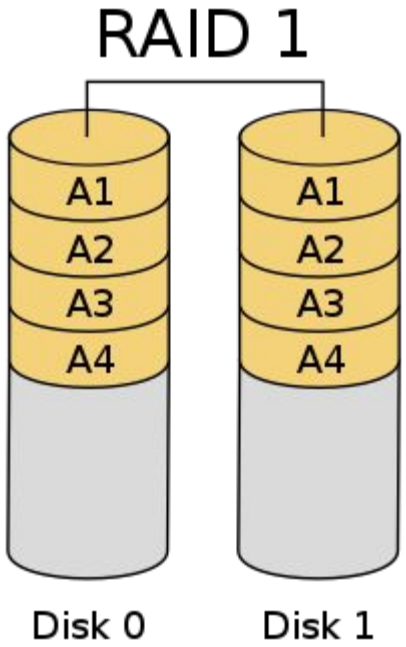


Zdroj obrázku:
https://en.wikipedia.org/wiki/Standard_RAID_levels#/media/File:RAID_0.svg

Disková pole RAID

RAID 1: spolehlivost

"mirroring"



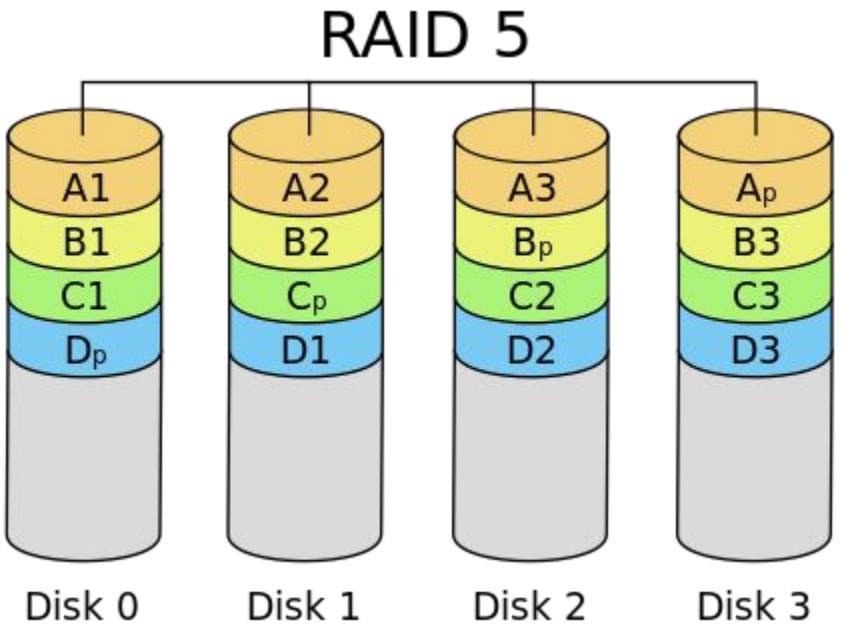
Zdroj obrázku:
https://en.wikipedia.org/wiki/Standard_RAID_levels#/media/File:RAID_1.svg

Disková pole RAID

RAID 5: spolehlivost a rychlost

"stripping + parity"

Přežije ztrátu jednoho disku

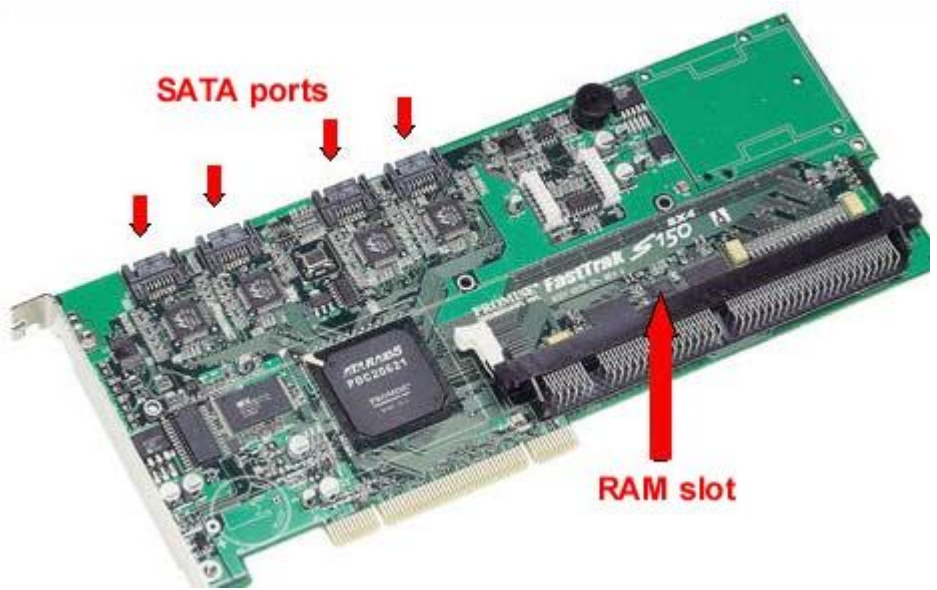


Zdroj obrázku:
https://en.wikipedia.org/wiki/Standard_RAID_levels#/media/File:RAID_5.svg

Disková pole RAID

HW pole = specializovaný Disk Array Controller

SW pole = běžné disky + SW kontrola



Disková pole RAID

Konfigurace RAID úrovní v Linuxu (mirror, stripe set apod.)

```
sudo apt install mdadm

sudo mdadm --examine /dev/sdb /dev/sdc

sudo mdadm --examine /dev/sdb1 /dev/sdc1

sudo mdadm --create /dev/md0 --level=mirror --raid-devices=2 /dev/sdb1 /dev/sdc1

cat /proc/mdstat

sudo mdadm --detail /dev/md0
```

Disková pole RAID

sudo mkfs.ext4 /dev/md0

sudo mkdir /mnt/raid1

sudo mount /dev/md0 /mnt/raid1

df -h /dev/md0

Disková pole RAID

Obnova po havárii disku

```
mdadm --manage /dev/mdN -r /dev/sdX1

mdadm --manage /dev/mdN -a /dev/sdY1
```

Logical Volume Manager (LVM)

- princip a vlastnosti
- návrh rozdělení disků v praxi
- správa LVM oddílů, vytváření, údržba
- přerozdělování místa mezi oddíly, přidání nového disku do LVM, zvětšní logického oddílu

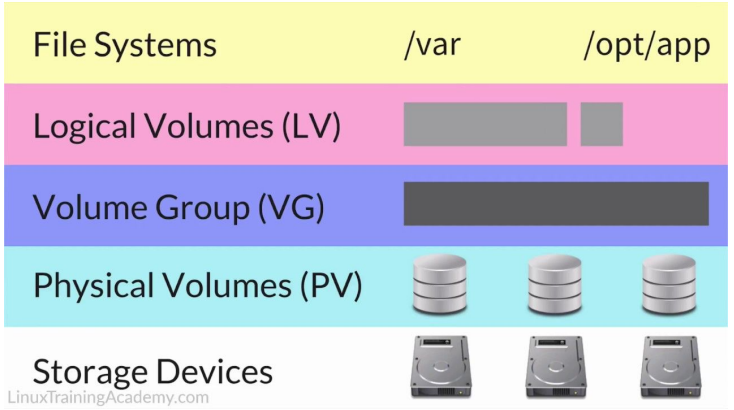
Logical Volume Manager (LVM)

Princip a vlastnosti:

"Virtualizace úložného prostoru"

lze vytvářet diskové oddíly sloučením a dělením, měnit jejich velikost za běhu a vytvářet snapshoty

Nástroje: `sudo apt install lvm2`



Logical Volume Manager (LVM)

Rozdělení disků v praxi, správa LVM oddílů, vytváření, údržba

Lze použít celý disk nebo jeho část (s LVM file systémem)

```
sudo pvcreate /dev/sda2  # Přidělíme úsek disku LVM

sudo vgcreate myVirtualGroup1 /dev/sda2  # Vytvoříme Volume group s tímto diskem

sudo vgextend myVirtualGroup1 /dev/sda3  # Rozšíříme Volume group
```

sudo lvcreate -n myLogicalVolume1 -L 10g myVirtualGroup1 # 10 GB virtual volume

Logical Volume Manager (LVM)

Rozdělení disků v praxi, správa LVM oddílů, vytváření, údržba

```
sudo mkfs -t ext4 /dev/myVirtualGroup1/myLogicalVolume1

sudo mkdir /test

sudo mount /dev/myVirtualGroup1/myLogicalVolume1 /test

df -h

sudo lvdisplay
```

Logical Volume Manager (LVM)

- **pvchange** — Change attributes of a Physical Volume.
- **pvck** — Check Physical Volume metadata.
- **pvcreate** — Initialize a disk or partition for use by LVM.
- **pvdisplay** — Display attributes of a Physical Volume.
- **pvmove** — Move Physical Extents.
- **pvremove** — Remove a Physical Volume.
- **pvresize** — Resize a disk or partition in use by LVM2.
- **pvs** — Report information about Physical Volumes.
- **pvscan** — Scan all disks for Physical Volumes.

Logical Volume Manager (LVM)

- **lvchange** — Change attributes of a Logical Volume.
- **lvconvert** — Convert a Logical Volume from linear to mirror or snapshot.
- **lvcreate** — Create a Logical Volume in an existing Volume Group.
- **lvdisplay** — Display the attributes of a Logical Volume.
- **lvextend** — Extend the size of a Logical Volume.
- **lvreduce** — Reduce the size of a Logical Volume.
- **lvremove** — Remove a Logical Volume.
- **lvrename** — Rename a Logical Volume.
- **lvresize** — Resize a Logical Volume.
- **lvs** — Report information about Logical Volumes.
- **lvscan** — Scan (all disks) for Logical Volumes.

Podrobný popis: https://wiki.debian.org/LVM

Nastavení ověřování přístupu uživatelů

- mechanismus fungování ověřování přístupu uživatelů
- konfigurace PAM
- použití externích ověřovacích metod

Nastavení ověřování přístupu uživatelů

Mechanismus fungování ověřování přístupu uživatelů

`/etc/passwd, /etc/shadow, /etc/group`

`passwd`

usermod

chmod, chgrp, chown

namespaces, cgroups

Nastavení ověřování přístupu uživatelů

PAM = Pluggable Authentication Modules

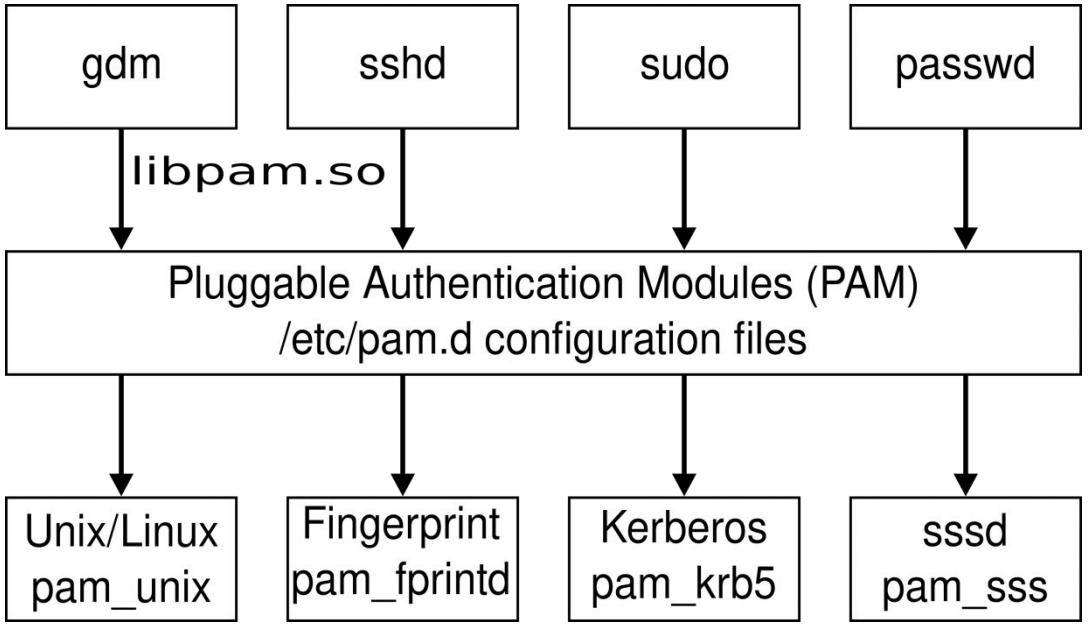
Centralizuje autentizaci a autorizace

`/etc/pam.d/*`

`/etc/security/*`

Zdroj obrázku:

https://www.redhat.com/sysadmin/pluggable-authentication-modules-pam



Nastavení ověřování přístupu uživatelů

Lze použít PAM pro řízení přístupu pro aplikaci sshd?

```
sudo ldd /usr/sbin/sshd | grep libpam.so
```

Příklad - zakázat root login přes ssh:

`/etc/pam.d/sshd`

```
auth    required        pam_listfile.so onerr=succeed \
item=user  sense=deny  file=/etc/ssh/deniedusers
```

(onerr chování, když soubor neexistuje)

Nastavení ověřování přístupu uživatelů

LDAP

Lightweight Directory Access Protocol

použitelný mj. pro centralizované ověřování

```
dn: cn=John Doe,dc=example,dc=com
cn: John Doe
givenName: John
sn: Doe
telephoneNumber: +1 888 555 6789
telephoneNumber: +1 888 555 1232
mail: john@example.com
manager: cn=Barbara Doe,dc=example,dc=com
objectClass: inetOrgPerson
objectClass: organizationalPerson
objectClass: person
objectClass: top
```

PAM + LDAP

`https://wiki.debian.org/LDAP/PAM`

Rekapitulace

- Přehled a troubleshooting startu systému
- Práce s disky
- Zálohování systému
- Sledování událostí systému
- Sledování vytížení systému
- Disková pole RAID
- Logical Volume Manager (LVM)
- Nastavení ověřování přístupu uživatelů

Dotazy?