# 12. Zálohování dat

* zajišťuje ochranu proti:
  + ztrátě dat díky hardwarové poruše
  + ztrátě dat díky uživatelské „blbosti“
  + poškození dat v důsledku různých úprav systému
  + výpadku části síťové infrastruktury
* nezajišťuje
  + dlouhodobé uložení dat (to řeší archivace)
* 2 základní podoby:
  + Onsite
    - uložení záloh v místě, jedná se především operační zálohy vhodné pro rychlou obnovu
  + Offsite
    - uložení záloh mimo budovu, nebo místnost se servery. Důvod Offsite zálohy je prostý, ochrana proti živelným katastrofám. Vhodné použít pro dlouhodobější ukládání (obvykle se offsite zálohy provádějí týdně, nebo měsíčně)
    - disaster recovery (DR) – soubor postupů, pro obnovu dat po ztrátě

# Proč vlastně zálohovat data?

* data jsou cennější než zařízení
* se zálohou v zádech je každá práce na počítačích hned veselejší
* uživatelé se budou se zálohou dat cítit daleko jistěji
* obnovení dat do původního stavu z existující zálohy je vždy rychlejší, než vytvoření dat do původního stavu

Zálohovací pojmy

* full backup – základní záloha, z ní se odvíjí ostatní typy záloh, trvá nejdéle, ale je kompletní
* differential, Incramental – podřízené zálohy, jsou malé, trvají krátce, ale nejsou kompletní; zálohují pouze změněné soubory od poslední full zálohy
* archive bit – vodítko, podle kterého se rozeznává, zda soubor byl nebo nebyl zálohován, dnes nahrazován tzv. journal systémem. Je to souborový atribut (podobně jako soubor pro čtení atd.) Hlavní nevýhodou je to, že znovu archivuje celý soubor jen po pouhém přejmenování.
* souborová záloha – zálohují se jednotlivé soubory jejich jednotlivým čtením a zápisem do záložního zařízení
* bitová záloha – nezálohuje se po souborech, ale po bitech. Je dramaticky rychlejší než souborová záloha (200 GB souborově může trvat cca 5-6 hodin, bitově cca 40 minut)
* VSS – volume shadow copy – technologie umožňující zálohu otevřených souborů; Windows only. Velmi podstatná pro zálohování.
* deduplikace – speciální algoritmus ukládání souborů, kdy se ukládají jen rozdílné bity, dochází k masivní úspoře diskového prostoru (pouze Windows 8 a server 2012)

# Zálohovací systémy pro domácnosti

# Obvykle se zálohuje ručně

# Nejčastější metodou jsou:

# Ruční kopie na externí médium (pevný disk, flash disk)

# Vypálení dat na optické médium

# Automatizované zálohování nebývá obvyklé.

# Zálohuje se výlučně systémem Full, nebo Copy.

# Typy zařízení:

# Externí disk – největší kapacita, nejrychlejší, nejjednodušší, náchylnější na poškození, přenášení možné, ale riskantní.

# Flash disk – uzpůsobený na přenášení, menší kapacity, přijatelná rychlost, ALE malá trvanlivost záznamů.

# Optické disky

# CD - na malé soubory, levné

# DVD - na běžná data, slušná výdrž, vyšší rychlost, levné

# BluRay – na větší data, slušná výdrž, vyšší rychlost, dost drahé

# Zálohovací systémy D2D (Disk to Disk)

# Nejjednodušší zálohování

# Jedná se o zařízení určené pro rychlé zálohování.

# V podstatě se jedná o „klasický malý“ server vybavený běžnými pevnými disky v RAID poli

# Obvykle je připojeno přes Ethernet rozhraní, nebo IEE1394, či USB

# Na vyšší úrovni se pak používá rozhraní iSCSI

# D2D si může dovolit každý, za D2D řešení lze považovat i externí disk

# D2D není vhodné pro archivaci

# Teoreticky nelimitovaná maximální kapacita

* Druhy
  + Interní pevné disky
  + Externí pevné disky - eSATA, RJ45, USB, FW
  + DAS RAID systémy - (přímo připojené diskové zálohovací systémy přes SCSI, nebo SAS)
  + NAS RAID systémy (iSCSI, RJ45) – pro menší firmy
  + D2D pro SAN (iSCSI) – dnes doménou větších firem
  + D2D pro SAN (FC) – pro největší firmy, obrovské kapacity
* Obvyklé jsou SATA disky (nebo Nearline, nebo MDL SAS disky)
* Obvykle se používá RAID5, nebo RAID6
* Používají se obvykle 3,5“ disky velkých kapacit
* Součástí D2D zařízení jsou obvykle deduplikační technologie
* Jedním z trendů D2D zařízení je identický formát s páskovými jednotkami, respektive jejich emulace, oproti zálohovacím systémům se chová jako páskové řešení

# Zálohovací systémy – pásková zařízení

* doménou větších firem (z českého pohledu)
* relativně drahé řešení
* vysoce spolehlivé řešení
* umožňuje snadné offsite zálohování
* dlouhodobě používané řešení s vysokou podporou od výrobců hardware i software
* největší dostupná kapacita na jedno zařízení (médium)
* teoreticky nelimitovaná kapacita
* omezený počet zápisů
* Druhy (dle médií)‏
  + DAT – nejstarší z aktuálně používaných technologií, vhodná pro malé firmy s malým objemem dat. Max. kapacita jedné pásky dnes 72GB
  + DLT – dnes již opouštěná pásková platforma, rychlá, vhodná pro větší firmy, max. kapacita na pásku 40/80GB
  + SDLT – jedna z enterprise technologií, extrémně rychlá, max. kapacita na pásku 320/640GB
  + LTO (Ultrium) – v současnosti nejoblíbenější, vhodné pro nejširší použití.
    - LTO1 – 100/200GB
    - LTO2 – 200/400GB
    - LTO3 – 400/800GB
    - LTO4 – 800/1600GB - umožňuje mimo jiné šifrování dat
    - LTO5 – 1600/3200GB
    - LTO6 – 3200/6400GB – ve vývoji
* Způsob zálohování na pásky:
  + Každá páska má dnes identifikaci v podobě Barcode (čárového kódu)
  + U pásky je klíčová kapacita
  + Na pásku se zapisuje v sekvenci, čili pokud chci zapisovat, musím projet vždy celé médium od začátku do konce.
  + Páska umožňuje i vícenásobný zápis do jednoho místa pásky (concurrent sessions)
  + Po zahájení zálohy se páska přetočí a ukládá postupně bity, po skončení zálohy se přetočí opět na začátek
  + Stejně funguje i obnova. Páska se musí vždy celá nejdříve přečíst, aby se provedla inventarizace katalogu uložených dat.
  + Páska má svoji dobu použitelnosti, obvykle cca 100 zápisových cyklů, pak dochází k expiraci.
  + Data však na pásce vydrží i mnoho let
  + Mechanika se musí pravidelně čistit pomocí čistícího média
* Výrobci: HP, Dell, IBM, Tanberg, Overland
* Druhy dle formátu zařízení – pásková zař.
  + Standalone mechaniky (samostatné mechaniky, tj. něco jako CD, DVD apod.)‏
    - Externí
    - Interní provedení
* Autoloadery – malé páskové roboty, 1-2 mechaniky s automatickým podávacím systémem, na 8-24 pásek
* Páskové knihovny – enterprise řešení, obrovský počet pásek (od 24-několik tisíc), větší množství mechanik, automatizované řídící systémy

## Zálohovací systémy - zálohovací software

* Pro domácnosti, malé firmy – Nero BackitUp, Acronis TrueImage, Symantec Ghost, Windows Backup
* Pro firmy
  + Obsluhují jak pásky, tak D2D
  + Obsahují obvykle software deduplikaci
  + Umějí zálohovat přes tzv. agenty (software komponenty v zálohovaném systému) při vysoké rychlosti
  + Umějí tzv. živé zálohy systému (ve Windows přes VSS, v Linuxu pomocí speciální technologie)
  + Podporují DR techniky
  + Umějí zálohovat bitově přímo diskové pole
  + Podporují zálohování virtualizačních řešení
  + =a=
  + Jsou velmi drahé
* Příklady pro firmy: HP Data Protector, IBM Tivoli, CA ArcServe, MS System Center Data Protection Manager, Veeam Backup pro Vmware

# Archivace

* slouží pro dlouhodobé uložení dat (ca více jak rok)
* soubory se nekopírují, ale přesouvají
* provádí se obvykle z důvodu ukládání dat mimo hlavního úložiště (šetří místo) – data se odstraní z hlavního úložiště a přesunou se do archivu, stejně jako u klasických papírových dokumentů
* archivace není vhodná místo zálohování, nezálohuje se většinou systém
* slouží pro dohledání určitých údajů po mnoha letech
  + v některých případech je nařízena zákonem (zdravotnictví, mzdový list, etc.)
* Druhy
  + CD, DVD mechaniky – trvalý záznam, relativně vysoká odolnost proti poškození při vhodném ukládání, nejlevnější forma zálohy
  + Páskové mechaniky – při použití speciálních pásek uložení až na 5-10 let
  + Celé systémy – udržovaný kompletní počítač (server) určený pro archivovaná data