

12 Zálohování dat

Obsah hodiny



Obsahem této hodiny je zabezpečení dat zálohováním a archivace dat.

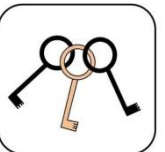
Cíl hodiny



Po prostudování budete schopni:

- vysvětlit důležitost zálohování, archivace dat
- vysvětlit rozdíl mezi zálohováním doma a ve firmě
- vybrat správné zálohovací médium
- zvolit správný typ zálohy
- charakterizovat jednoduché a víceúrovňové zálohování
- orientovat se v různých možnostech zálohování

Klíčová slova



Archivace, Zálohování, Zálohovací média, Úplná, rozdílová a přírůstková záloha, Image, RAW image, Deduplikace, RAID, Zrcadlení

Ztrátě dat lze sice předcházet, ale ne zabránit. A tak je nutné se alespoň snažit minimalizovat následky. Jak? Jednoduše pečlivým a pravidelným data zálohováním.

12.1 Archivace versus zálohování

Záloha se vytváří z důvodu ochrany dat, tak aby v kterémkoli okamžiku byly k dispozici kopie dat pro co nejrychlejší obnovu pokud možno v aktuálním stavu.

Důvodem archivace je zachování určitého stavu dat a zajištění jejich dostupnosti z dlouhodobého hlediska. Problematiku archivace podrobně řeší Zákon České národní rady č. 97/1974 Sb. ze dne 17. října 1974 o archivnictví ve znění Zákona České národní rady č. 343/1992 Sb.

12.2 Záložní a archivační média

Před samotným zálohováním je potřeba vybrat média pro zálohování a archivaci dat. K dispozici je řada možností:

- výměnná média
 - diskety
 - optické disky
 - CD (700 MB), DVD (4,7 - 17,1GB), HD DVD (15-60 G), Blu-ray disk 25-80 GB
 - magnetické, magnetooptické pásky - levnější varianta než disky, kapacita je na úrovni pevných disků ... až 1TB
 - streamery (pásková zařízení)
 - DAT - modernější obdoba streamerů
 - výměnné pevné disky
 - WORM disk (Write Once, Read Many, magneto-optický disk, 128 MB až po 600 GB)
- nevýměnná média
 - další disk
 - disk mirroring
Mirroring je možno využít pro disky, které jsou zapojeny v RAID1. Mirroring probíhá tak, že data jsou v reálném čase zapisována současně na dva pevné disky.
 - clusterování, duplexing

Je vždy vhodné zvážit řadu okolností, které ovlivní výběr vhodného záložního média. Například: jaké množství dat se bude zálohovat (dostatečná kapacita média), nakolik jsou pro data důležitá (kolik do záložních médií investovat), jak dlouho se bude nutné data uchovat (záložní médium musí být i po letech čitelné), jak často se bude záloha provádět, zda a jak často se bude záložní médium přepisovat (kolikrát lze médium přepsat, aby bylo ještě použitelné), kdo bude zálohování provádět (míra složitosti při zápisu na médium).

Výběr média pro zálohování se tedy provádí s ohledem na jeho

- kapacitu,
- rychlost,
- trvanlivost (životnost),
- dostupnost,
- použitelnost,
- složitost,
- spolehlivost,
- cenu.

Specifickou vlastností zálohovacích médií je jejich životnost. Životnost média zkracuje v podstatě každé jeho použití. Je nutná opatrnost při jakékoli manipulaci s ním. Existuje celá řada faktorů, které zkracují životnost médií, jsou to:

- vlhkost,
- prašnost,
- světelné záření,
- kouř a další.

Například výrobce CD a DVD Verbatim uvádí, že životnost jejich médií za ideálních podmínek je až 100 let. Tyto údaje je třeba brát s rezervou – ne vždy lze zajistit právě ty ideální podmínky.

Pro zálohování doma se často používá pevný disk, což je způsob nejrychlejší, ale nejméně bezpečný zejména, když u varianty, kdy budou data i záloha na stejném disku a na stejném diskovém oddílu. Vhodnější je uchovávat zálohu dat na disku externím nebo síťovém.

Nejčastějším vyměnitelným (externím) médiem pro uchování dat doma, či v malých firmách jsou CD a DVD disky. Výhodou je cena, dostupnost, což umožňuje pořídit více záložních kopií. Nevýhodou je pracnost aktualizace.

Ve firmách se běžně používá zálohování dat na síťová úložiště (sdílený disk, ftp server, ...). Jednotliví uživatelé ukládají svá pracovní data do sdíleného úložiště na serveru nebo síťových discích. Za uchování dat ve sdílených úložištích pak obvykle odpovídá správce sítě, který tato data dále zálohuje např. na pásku.

Zálohování dat na pásku je obvyklé ve větších organizacích. Zálohy na pásku se ukládají prostřednictvím páskových mechanik nebo prostřednictvím zálohovacích knihoven (autoladerů).

V současné době lze provádět zálohování dat do virtuálního prostředí, což je metoda vhodná pro zálohování funkčního systému, kterou je možno provozovat ve virtuálním prostředí v případě havárie samotného fyzického stroje.

12.3 Typy záloh

Nejjednodušší způsob, jak zálohovat data je jejich prosté zkopírování do záložního úložiště. Pokud je dat větší množství, je možné data komprimovat.

- MS Windows: Winzip a Winrar
- Linux: zip/gzip

Pozor, při pokud při komprimaci dojde k poškození dat, nebude možno zálohu obnovit. Tento způsob je vhodný u pro zálohování jednotlivých souborů nebo souborů jednotlivých uživatelů.

Pokud je třeba zálohovat a archivovat data více uživatelů, data na síťových discích, pak se pro zálohování využívají speciální nástroje, které umožňují podle potřeby vytvářet různé typy záloh včetně záloh komprimovaných a zjednodušují následnou obnovu dat.

Obecně se rozlišují tři základní typy záloh:

- **úplná:** (Full backup nebo Partial backup) zálohuje se vše nebo se provádí plná záloha jen vybraných dat nebo datových oblastí
- **rozdílová** (diferenciální) záloha: zaznamenává změny proti poslední úplné záloze
- **přírůstková** (inkrementální) záloha: zaznamenává změny proti poslední přírůstkové záloze.

Pokud se např. v pondělí vytváří úplná záloha a každý den do pátku zálohy rozdílové, bude každý den k dispozici záloha všeho, co se změnilo od pondělní úplné zálohy:

- v úterý vše, co se vytvořilo v úterý,
- ve středu vše co se vytvořilo v úterý a ve středu,
- ve čtvrtek vše co se vytvořilo v úterý, ve středu, ve čtvrtek
- v pátek vše co se vytvořilo v úterý, ve středu, ve čtvrtek a v pátek

Pokud se např. v pondělí vytvoří úplná záloha a každý den do pátku zálohy přírůstkové, budou denní zálohy obsahovat to, co se změnilo od posledního zálohování.

Z pohledu toho kdy zálohu provádíme (plánování zálohování) rozdělujeme zálohy na

- off-line backup (záloha v době mimo provoz),
- on-line backup (průběžná záloha) dochází k zálohování za plného provozu systému.

12.4 Metody zálohování

Jednoduché zálohování

Data se zálohují každý den v týdenních cyklech. Tato metoda má dvě úrovně:

- **Úplné** zálohování (angl. full backup), kterým se vytvoří zálohy všech dat na začátku cyklu.
- Denně se provádějí **přírůstkové** zálohování.

Př.: data zálohujeme každý den, k dispozici je šest páskových kazet:

- *první pásku použijeme (např. v pátek) pro uložení první úplné zálohy*
- *pásy 2 až 5 použijeme pro přírůstkové zálohování (od pondělí do čtvrtka)*
- *pak vytvoříme novou úplnou zálohu na pásku číslo 6 (druhý pátek) a začnete znovu s inkrementálním zálohováním na pásy 2 až 5.*
- *pásku číslo jedna nepřepisujeme do doby, než se ukončí nové úplné zálohování - v jeho průběhu by se totiž mohlo stát něco neočekávaného.*
- *poté, co vytvoříme novou úplnou zálohu na pásku číslo 6, pásku 1 uschováme někde jinam pro případ, že by se ostatní zálohovací pásy zničily, např. při požáru.*
- *když pak potřebujete vytvořit další úplnou zálohu, dojdete pro pásku číslo 1 a pásku číslo 6 necháte na jejím místě.*

Víceúrovňové zálohování

Víceúrovňovým zálohováním lze historii vytváření archivu prodloužit o další časový horizont. Pořizují se zálohy:

- měsíční,
- týdenní
- denní

Př.: S deseti páskovými kazetami, můžeme prodloužit historii zálohování ze dvou týdnů (s využitím všech denních záloh) na dva měsíce:

- *pásy 1 a 2 použijeme na měsíční zálohy (první pátek každého měsíce)*
- *pásy 3, 4, 5 a 6 na týdenní zálohy (všechny ostatní pátky)*
- *pásy 7, 8, 9 a 10 na denní zálohy (od pondělí do čtvrtka).*

12.5 Uložení záložních kopií

Význačné záložní kopie se uchovávají na vzdálených bezpečných místech. Jednak z důvodu zabezpečení proti odcizení, jednak jako ochrana před případnou katastrofou.

Běžné denní zálohy další ukládáme blízko, protože je potřebujeme mít po ruce pro případ rychlé obnovy dat v důsledku smazání dat či jiných menších "katastrof".

12.6 Co zálohujeme

Při zálohování je třeba zvážit, která data zálohovat a na základě důležitosti a toho jak se obsah souborů v čase mění pak vybírat metodu a typ zálohování.

Zálohují se zejména:

- uživatelské soubory v domovských adresářích
- data aplikací
- systémové konfigurační soubory
- databáze.
- poštovní schránky,
- log-soubory

Nemusí se zálohovat software, jenž lze jednoduše obnovit z instalačních médií nebo virtuální soubory jako je např. v Linuxu souborový systém /proc, který se vytváří při každém spuštění systému. (zejména soubor /proc/kcore, což je aktuální obraz fyzické paměti).

12.7 Další možnosti zálohování

Image disku

Neboli obraz pevného disku je záloha, která zálohuje veškerý obsah pevného disku včetně informací o zavádění operačního systému a záznamu Master Boot Record. Metody tvorby obrazů pevných disků bývají obvykle postaveny na technologii zálohování sektorů na disku. Dle použitého software jsou buďto zálohovány jen obsazené sektory na disku nebo jsou zálohovány všechny sektory na disku nezávisle na tom, zda jsou obsazené (RAW image). V obou případech je výsledkem zálohování disku záložní soubor. Díky této záloze je možné do výchozího stavu obnovit buďto původní disk, nebo může být tato záloha obnovena na disk jiný.

RAW image – bitová kopie disku

Je záloha typu obraz disku vytvořená zálohováním veškerých sektorů na disku. RAW image neumožňuje využití komprese, takže výsledná záloha je stejně velká, nebo o málo větší než takto zálohovaný disk. RAW image jsou hojně využívány při záchraně dat, kdy se vytvoří přesný „otisk“ pevného disku a požadovaná data jsou obnovována z této kopie původního disku.

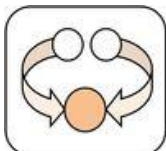
Deduplikace

Zálohování s deduplikací je takové zálohování, kdy jsou vyskytující se redundantní data ukládána do zálohy pouze jednou. Deduplikace může podstatně snížit nároky na velikost úložiště. Zdroje dodavatelů softwarových řešení pro zálohování s deduplikací uvádějí zmenšení záloh až o 90%. V závislosti na použitém řešení je možno deduplikaci využít jak pro jednotlivé počítače, tak pro celou síť počítačů.

RAID

RAID neboli Redundant Array of Independent Disks je typ diskových řadičů, který umožňuje koordinovanou práci dvou a více pevných disků. Pevné disky zapojené v RAIDu tvoří tzv. RAIDové pole. Existuje několik typů raidových polí. Pro ochranu před hardwarovou poruchou pevného disku je často využíván RAID1 neboli zrcadlení disků - data jsou v reálném čase zapisována současně na dva pevné disky.

Shrnutí kapitoly



Zálohování se provádí z důvodu

- minimalizace následků při ztrátě nebo poškození dat
- archivace dat (často vyžadovaná zákonem)

Podle řady kritérií (kapacita, rychlost, trvanlivost a životnost, dostupnost, použitelnost, složitost zápisu, spolehlivost, cena) a s ohledem na množství dat je třeba vybrat médium pro zálohování a zvolit si metodu zálohování (jednoduché, víceúrovňové zálohování).

Pro realizaci zálohování se pak používají tři typy záloh:

- Úplná
- Rozdíllová
- Přírůstková

Kopie je nutné ukládat na bezpečné místo tak, aby byly k dispozici pro rychlou obnovu dat, některé kopie naopak na místa vzdálená pro případ katastrofy.

Důležité je správně posoudit, která data a jak často zálohovat.

Dále je možno pro zálohování dat využít:

- Image disku
- RAW image – bitová kopie disku
- Deduplikace
- Zrcadlení disků

Kontrolní otázky a úkoly



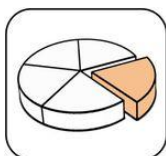
- 1) Jaké jsou důvody pro zálohování dat?
- 2) Jmenujte kritéria pro výběr zálohovacího média.
- 3) Co ovlivňuje životnost zálohovacího média?
- 4) Porovnejte zálohu rozdílovou a přírůstkovou.
- 5) Vysvětlete princip jednoduchého zálohování.
- 6) Jak se liší jednoduché a víceúrovňové zálohování?
- 7) Co nemusíme zálohovat?

Otázky k zamyšlení



- 1) Prozkoumejte, které nástroje se pro zálohování používají v OS MS Windows a Linux
- 2) Zjistěte, jaké volně šiřitelné programy lze použít pro zálohování.
- 3) Která data je třeba zálohovat v co nejkratším časovém horizontu a naopak?
- 4) Která data budou vyžadovat velkou frekvenci zálohování?
- 5) Co jsou to online zálohovací služby, jaké jsou jejich výhody a nevýhody?

Použitá literatura a jiné zdroje:



- [1] PŘIBYL, Tomáš. Zálohování a archivace dat z paměťových médií . ICT Security [online]. 03.08. 2009, 1, [cit. 2011-11-16]. Dostupný z WWW: <<http://www.ictsecurity.cz/odborne-clanky/zalohovani-a-archivace-dat-z-pametovych-medii.html>>.
- [2] Linux documentation Project (CS) : Průvodce správce operačního systému Linux [online]. Praha : Computer Press, 16.01.2008 [cit. 2011-11-16]. Dostupné z WWW: <<http://knihy.cpress.cz/linux-d5.html#>>, .nebo <<http://www.nuc.elf.stuba.sk/lit/ldp/02/020-10.htm>>.
- [3] ACRONIS STORE.EU. Zálohování.net [online]. [cit. 2011-11-16]. Dostupné z WWW: <<http://www.zalohovani.net/>>.