

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra informačního inženýrství



Bakalářská práce

Zálohování relačně databázových evidencí

Miroslav Smutný

© 2017 ČZU v Praze

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Miroslav Smutný

Informatika

Název práce

Zálohování relačně databázových evidencí

Název anglicky

Backup of relational databases

Cíle práce

Bakalářská práce je zaměřena na problematiku zálohování databázových evidencí. Hlavním cílem této práce je:

- a) objasnit teoretické principy relačně databázové technologie se zřetelem na problematiku zálohování a obnovy databázových evidencí,
- b) zmapovat momentální stav této problematiky a vymezit její relevantnost včetně požadavků na ni kladených,
- c) navrhnout přijatelné řešení této problematiky v souladu s identifikovanými požadavky,
- d) ověřit funkčnost navržených záležitostí,
- e) ověřené záležitosti zobecnit pro další možná uplatnění.

Závazný harmonogram:

Teoretické principy, literární rešerše – předmět 1. zápočtu z BP: do 5.9.2016

Zmapování momentální situace řešené problematiky, identifikace požadavků na ni kladených – do 31.11.2016

Navržení možného řešení se zřetelem na identifikované požadavky – předmět 2. zápočtu z BP: do 10.1.2017

Ověření a zobecnění navrhovaných záležitostí – předmět 3. zápočtu z BP: do 10.3.2017

Metodika

Použitá metodika zadané bakalářské práce bude založena na studiu a analýze dostupných informačních zdrojů, zkušeností jejího autora a existujících řešení v dané oblasti. Stěžejní pro vypracování této závěrečné práce budou metody a techniky relačně databázové technologie v kontextu s touto problematikou. Navrhované řešení bude zohledňovat identifikované požadavky a očekávání spojená s řešenou záležitostí. Na podkladě syntézy teoretických poznatků a dosažených výsledků budou formulovány závěry této bakalářské práce a následně zobecněny pro další možná použití.

Doporučený rozsah práce

45-55 stran

Klíčová slova

relačně databázová evidence, zálohování dat, SQL, databázové transakce

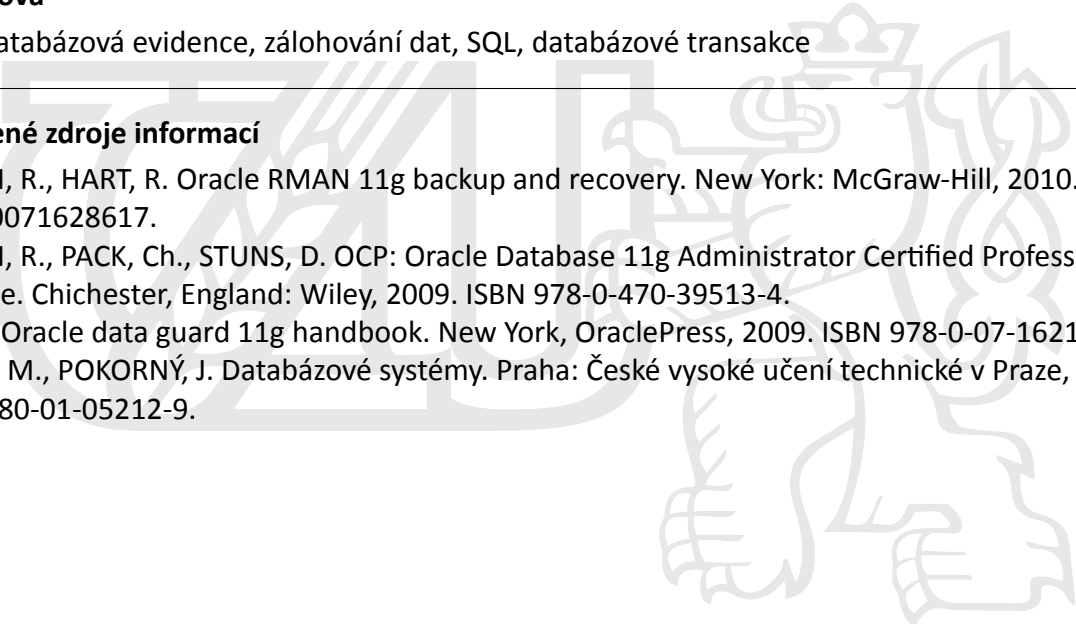
Doporučené zdroje informací

FREEMAN, R., HART, R. Oracle RMAN 11g backup and recovery. New York: McGraw-Hill, 2010. ISBN 9780071628617.

FREEMAN, R., PACK, Ch., STUNS, D. OCP: Oracle Database 11g Administrator Certified Professional Study Guide. Chichester, England: Wiley, 2009. ISBN 978-0-470-39513-4.

LARRY, C. Oracle data guard 11g handbook. New York, OraclePress, 2009. ISBN 978-0-07-162148-9.

VALENTA, M., POKORNÝ, J. Databázové systémy. Praha: České vysoké učení technické v Praze, 2013. ISBN 978-80-01-05212-9.



Předběžný termín obhajoby

2016/17 LS – PEF

Vedoucí práce

doc. Dr. Ing. Václav Vostrovský

Garantující pracoviště

Katedra informačního inženýrství

Elektronicky schváleno dne 21. 2. 2017

Ing. Martin Pelikán, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 21. 2. 2017

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 13. 03. 2017

Čestné prohlášení

Prohlašuji, že svou bakalářskou práci "Zálohování relačně databázových evidencí" jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu použitých zdrojů na konci práce. Jako autor uvedené bakalářské práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 15.3.2017

Poděkování

Rád bych touto cestou poděkoval doc. Dr. Ing. Václavu Vostrovskému za odborné vedení této práce a veškerý věnovaný čas. Dále bych chtěl poděkovat společnosti Your System, spol. s.r.o. za prostory poskytnuté v době vykonávání bakalářské praxe, zejména jejímu zaměstnanci Ing. Vladimíru Stejskalovi za jeho cenné rady, znalosti a čas, které mi věnoval v průběhu této praxe. Nakonec bych chtěl poděkovat své rodině a přátelům za podporu v průběhu celého studia a v průběhu zpracování této bakalářské práce.

Zálohování relačně databázových evidencí

Souhrn

Tato bakalářská práce se zabývá problematikou zálohování relačních databázových evidencí. Teoretická část popisuje typy souborů, které je nutné zálohovat a také související databázové systémy. Dále jsou v této části popsány požadavky na zálohu, principy samotného zálohování, zálohovací nástroje a také strategie zálohování a obnovy.

V praktické části je provedeno zmapování současné situace ve dvou společnostech a také analýzy náročnosti jednotlivých typů záloh a jednotlivých principů zálohování. Následně je navržena fiktivní společnost a jsou identifikovány její požadavky na strategii zálohování a obnovy. Principy zálohování jsou analyzovány pomocí vícekriteriální analýzy variant. Po výběru optimálního principu zálohování je strategie realizována na testovací databázi. V praktické části jsou také zobrazeny skripty pro vytváření záloh.

Nakonec je zhodnocena strategie zálohování a samotné řešení je zobecněno pro další využití v praxi.

Klíčová slova: relačně databázová evidence, zálohování dat, SQL, databázové transakce

Backup of relational databases

Summary

This bachelor thesis deals with relational database backup issues. Theoretical part describes types of files, which needs to be backed up and related database systems. Further, this section describes backup requirements, principles of backup itself, backup tools and backup and recovery strategy.

Practical section evaluates present situation of relational database backup in two companies and analyses demandingness of individual backup types and principles. Then, the fictional company is proposed and its backup and recovery strategy requirements are identified. Principles of backup are analysed with help of multi-criteria analysis. After choice of optimal backup principle, proposed backup and recovery strategy is implemented on test database. There are, in practical section, shown scripts for backup.

At the end, backup and recovery strategy is evaluated and solution itself is generalized for implementation.

Keywords: relational database records, data backup, SQL, database transactions

Obsah

1 Úvod.....	11
2 Cíl práce a metodika	13
2.1 Cíl práce	13
2.2 Metodika	13
3 Teoretická východiska	15
3.1 Soubory databázového systému	15
3.1.1 Datové soubory	15
3.1.2 Konfigurační soubory	16
3.1.3 Transakční záznamy	16
3.2 Databázové systémy	17
3.2.1 Oracle Database	17
3.2.2 Microsoft SQL Server	18
3.3 Zálohování a archivace	19
3.4 Rozdělení záloh.....	19
3.4.1 Komprimace.....	19
3.4.2 Plná a částečná záloha.....	20
3.4.3 Přírůstek	21
3.4.4 Dostupnost databáze v průběhu zálohování.....	21
3.5 Požadavky na zálohové soubory	21
3.5.1 Ochrana	21
3.5.2 Duplicita.....	23
3.5.3 Fyzické rozdělení záloh dat a záznamů	24
3.6 Validace záloh	24
3.7 Obnova	24
3.7.1 Politiky udržování záloh	24
3.8 Strategie zálohování a obnovy	25
3.9 Zálohovací nástroje	26
3.9.1 Recovery Manager – Oracle Database.....	26
3.9.2 Management Studio – Microsoft SQL Server	27
3.10 Principy zálohování a obnovy	27
3.10.1 Export a import	27
3.10.2 Manuální záloha a obnova	28
3.10.3 Specializovaný nástroj	28
4 Praktická část	29
4.1 Vliv datového nárůstu na velikost záloh	29

4.1.1	Plná záloha.....	30
4.1.2	Přírůstkové kumulativní zálohy	31
4.1.3	Zálohy archivních záznamů	31
4.2	Analýza současné situace	32
4.2.1	Charakteristika společností	32
4.2.2	Strategie zálohování.....	33
4.2.3	Zhodnocení strategií zálohování	33
4.3	Identifikace požadavků	34
4.4	Definice společnosti	34
4.4.1	Charakteristika společnosti	34
4.4.2	Požadavky na strategii zálohování.....	34
4.5	Návrh strategie zálohování.....	35
4.6	Analýza náročnosti.....	36
4.6.1	Export.....	36
4.6.2	Manuální záloha.....	37
4.6.3	Specializovaný nástroj	38
4.7	Výběr optimálního principu zálohování.....	39
4.8	Nastavení zálohování dle strategie.....	42
4.9	Zhodnocení řešení	44
4.10	Návrh vylepšení	45
5	Výsledky a diskuse	46
5.1	Zobecnění řešení	46
5.2	Diskuze.....	46
6	Závěr.....	48
7	Seznam použitých zdrojů	50
8	Přílohy	51

Seznam obrázků

Obrázek č. 1 - Vývoj velikosti souborů plné zálohy.....	30
Obrázek č. 2 - Vývoj velikosti souborů přírůstkové kumulativní zálohy.	31
Obrázek č. 3 – Vývoj velikosti souborů zálohy archivních záznamů.....	32
Obrázek č. 4 – Graf časové náročnosti exportu	37
Obrázek č. 5 – Graf časové náročnosti manuální zálohy.....	38
Obrázek č. 6 – Graf časové náročnosti RMAN	39

Seznam tabulek

Tabulka č. 1 – Denní frekvence zálohování	33
Tabulka č. 2 – Časová náročnost zálohování pomocí exportu	36
Tabulka č. 3 – Časová náročnost manuálního zálohování.....	37
Tabulka č. 4 – Časová náročnost zálohování pomocí RMAN.....	38
Tabulka č. 5 – Saatyho matice pro odvětví kritických systémů	40
Tabulka č. 6 – Saatyho matice pro odvětví běžných systémů	41
Tabulka č. 7 – Výpočet užitku pro kritické systémy	41
Tabulka č. 8 – Výpočet užitku pro běžné systémy	41
Tabulka č. 9 – Výsledné zhodnocení principů zálohování	41

1 Úvod

Bakalářská práce se zabývá tématem zálohování relačně databázových evidencí. Obecně je zálohování velmi potřebné a důležité. Z hlediska ztráty dat není rozdíl mezi záznamy v databázových evidencích a soubory, jako jsou textové dokumenty, multimédia a další, uloženými ve složkách. Pokud jsou data preventivně zálohována, není velký problém v případě ztráty nebo poruchy tato data obnovit. Tato problematika se v posledních letech velmi rozrostla z důvodu nárůstu objemu dat a velkého požadavku na rychlost zálohování těchto dat. Proto je potřeba vynaložit mnoho zdrojů na uložení a výkon pro zpracování a uložení kopií dat. Relační databázové evidence jsou hromadně využívány v podnicích díky vysoké rozšířenosti, rychlému zpracování dat, a velkému zájmu ze strany databázových administrátorů. Je nutností tyto relační databázové evidence kvalitně a spolehlivě zálohovat. Existuje několik společností vyvíjejících relační databázové systémy, obsahující vlastní zálohovací nástroje. Tyto nástroje mají mnoho společného, například principy a postupy zálohování. Terminologie se ovšem může lišit. Zálohování velkých databázových evidencí s vysokým množstvím přesunu dat je nutné zautomatizovat. Není v silách žádného administrátora provádět zálohování manuálně. Existují zálohovací strategie, které vyžadují zálohu záznamů o pohybu dat každých 15 minut i méně. Administrátor by tedy nedělal nic jiného, než zadával příkazy pro spuštění zálohování.

Podniky do relačních databázových evidencí ukládají citlivá interní i externí data. Jako interní data se rozumí informace o zaměstnancích a data strategického a taktického významu. Oproti tomu externí data jsou spjata se zákazníky, případně s konkurencí a trhem. Taková data je nutné efektivně zálohovat pro případné poruchy informační infrastruktury. Informace jsou pro podniky zdrojem konkurenceschopnosti a data, jež jsou nositeli informací je nutné správně a efektivně zálohovat pro případ poruchy nebo útoku konkurence.

Zálohování jako takové je požadováno v každém podniku, který provozuje relačně databázovou evidenci. V případě, že podnik nezálohuje svá data, se může velmi snadno vyskytnout problém z lidské strany, případně i v důsledku selhání jednotlivých prvků systému a tím dojít ke ztrátě dat.

Důležitost zálohování spočívá v uchování databázových informací a jejich případné rychlé obnově, díky čemuž může firma pokračovat ve své činnosti. Činnost některých firem je závislá z velké části na databázových evidencích a každé zdržení v činnosti by

znamenalou velkou ztrátu zisku nebo klientů a jejich přechod ke konkurenci, v extrémních případech i ekonomické problémy na straně klientů.

Zálohování je tedy užitečné z důvodu zachování konkurenceschopnosti a udržení klientů a tedy i zisku.

2 Cíl práce a metodika

2.1 Cíl práce

Bakalářské práce je zaměřena na problematiku zálohování relačně databázových evidencí a s tím spjatých procesů. Hlavním cílem této práce je navržení optimální strategie zálohování a obnovy na základě požadavků specifikovaných v teoretické části. Dílčími cíli je objasnit teoretická východiska a identifikovat požadavky se zřetelem na problematiku zálohování a obnovy databázových evidencí, zmapování momentální strategie zálohování ve vybrané firmě a vymezení její relevantnosti včetně požadavků na ní kladených, ověření funkčnosti a zhodnocení navrženého řešení. Následně je navržené řešení zobecněno pro další použití.

2.2 Metodika

Metodika zpracování této bakalářské práce je založena na studiu a analýze dostupných informačních zdrojů a existujících řešení v oblasti zálohování relačně databázových evidencí. Stěžejní pro vypracování budou metody a techniky relačně databázové technologie v kontextu s problematikou zálohování. Konkrétně se jedná o technologii transakčního zpracování. Navržené řešení bude zohledňovat identifikované požadavky a očekávání spojená s řešením. Na podkladě syntézy teoretických poznatků a dosažených výsledků bude řešení zobecněno pro další možná použití a budou formulovány závěry bakalářské práce.

Zvolená metodika pro realizaci praktické části této bakalářské práce spočívá v těchto krocích:

1. Analýza shromážděných zdrojů a jejich kritické zhodnocení.
 - Studium informačních zdrojů a seznámení s existujícími řešeními v dané oblasti.
2. Analýza náročnosti zálohovacích principů
 - Identifikace časové náročnosti zálohování s ohledem na velikost datových souborů a následných zálohových souborů. Zhodnocení principů zálohování použitím vědeckých metod. Konkrétně se jedná o použití metody vícekritériální analýzy variant.

3. Návrh strategie zálohování a obnovy
 - Charakteristika fiktivní společnosti, jejích požadavků a následný návrh strategie zálohování s využitím teoretických východisek.
4. Nastavení a otestování funkce zálohování na základě navržené strategie
 - Nastavení zálohování dle navržené strategie zálohování použitím identifikovaného optimálního principu.
5. Prověření výsledků testování a optimálnosti navržené strategie.
 - Posouzení výsledků zálohování a obnovy a následné prověření funkce samotné funkcionality navržené strategie. Kontrola navržené strategie a výsledků testování, případný návrat k návrhu strategie zálohování a obnovy.
6. Implementace navrženého řešení.
 - Navržené řešení je nastaveno na požadované databázi.

3 Teoretická východiska

3.1 Soubory databázového systému

Soubory, popsané v této kapitole, je třeba zálohovat pro zachování konzistence databáze. Tyto soubory slouží k uchování dat nebo záznamů o operacích v databázi.

3.1.1 Datové soubory

Datové soubory jsou potřebné k dlouhodobému ukládání dat. Jejich vnitřní struktura může být jednoduchá (jedna tabulka) nebo složitá (více tabulek). Každý databázový systém používá jiný způsob ukládání. Data uložená v těchto souborech musí být uložena fyzicky na disku. Termín „**datafile**“ používá většina databázových systémů. Z hlediska ukládání dat se databáze liší způsobem logického ukládání dat. Databázový systém zprostředkovává uživatelům přístup k datům, ale jen z logického hlediska. Uživatelé stačí zobrazit pouze své databázové objekty, o samotné ukládání dat se nestarají. (1, s. 1-8)

Dalším důležitým pojmem je „**tablespace**“ případně „**filegroup**“. Tyto pojmy značí logickou strukturu databáze. V obou případech je princip stejný a to logicky sdružovat související data. Jedna logická struktura může obsahovat více datových souborů. Jeden datový soubor ovšem nemůže být přiřazen k více logickým strukturám. Díky více datovým souborům v logické struktuře je možné rozdělit zátěž při čtení nebo zápisu dat na více uložišť, případně i jiné počítačové systémy. Tímto způsobem však může být ohrožena konzistence dat. Logické struktury mohou být v režimu online nebo offline. V případě, že je nutné upravit nastavení logické struktury, je přepnuta do režimu offline a tím je zabráněno uživatelům v přístupu k datům. Po úpravě nastavení a pro zpřístupnění dat je logická struktura přepnuta zpět do režimu online. (2, s. 309)

Z hlediska zálohování je jediný rozdíl mezi logickou a fyzickou strukturou takový, že pokud je zálohována logická struktura, jsou automaticky vybrány všechny datové soubory, které jsou ke struktuře přiřazeny. V případě zálohování fyzické struktury je možné vybrat pouze určité datové soubory. (1, s. 15-4)

3.1.2 Konfigurační soubory

Prvním typem jsou kontrolní soubory, které obsahují název databáze, časové razítko vytvoření databáze, popis fyzické struktury, umístění datových souborů, transakčních záznamů a zálohovaných transakčních záznamů. Kontrolní soubory jsou velmi důležité pro zálohování a obnovu, a proto je výhodné udržovat více jejich záloh. (3, s. 64)

Druhým typem jsou konfigurační soubory, které jsou využívány při spouštění databáze a obsahují informace o velikosti přidělené paměti databázi, název databáze, co provádět s naplněnými transakčními záznamy, název a umístění kontrolních souborů a názvy logických struktur pro odvolání transakcí. Konfigurační soubory jsou klíčové pro spuštění databáze, a proto je nutné udržovat více záloh těchto souborů. (1, s. 12-2)

3.1.3 Transakční záznamy

Transakce je skupina SQL příkazů, které je nutné provést jako celek, nebo případně vrátit změny dat provedené těmito příkazy, pokud se v průběhu provádění příkazů stal nějaký incident. Transakce je logická, atomická jednotka práce. Velmi dobrým příkladem je převod peněz z účtu na jiný účet. Tento převod je možné rozdělit na 3 úkony. Prvním je snížení zůstatku na účtu odesílatele. Další je zvýšení zůstatku na účtu příjemce. Posledním úkonem je uložení transakce do záznamů. V případě potřeby jsou zůstatky na účtech navráceny do původních hodnot. Tím je zachována konzistence dat. (4, s. 1-6)

Transakční záznamy je nutné uchovávat z důvodu udržení konzistence dat. Ty se uchovávají v určených souborech. Pokud se transakce z určitého důvodu nedokončí, díky těmto záznamům je možné vrátit databázi do stavu před započítáním transakce a díky tomu zachovat konzistenci dat. V Oracle Database se tyto soubory nazývají „**redo log**“, v Microsoft SQL Server „**transaction log**“. Zálohy těchto záznamů slouží jako podpůrný prvek k celkovým zálohám. Není možné plně zálohovat větší databáze každý den a zálohy těchto záznamů mají uloženou historii transakcí, většinou od plné nebo přírůstkové zálohy, které jsou při obnově provedeny. Díky tomu je možné obnovit databázi do nejbližší doby k incidentu. (2, s. 311) (3, s. 62)

3.2 Databázové systémy

Databázový systém je software, obsahující systém řízení báze dat a samotnou bázi dat. Má na starost organizaci, ukládání a načítání dat, případně metadat. Systém řízení báze dat umožňuje přístup k datům a jejich manipulaci. Báze dat jsou pouze data, specificky uložena a načítána právě systémem řízení. Databázové systémy mají několik prvků. Prvním z nich je samotné jádro, které se stará o paměť a uložení. Dalším prvkem bývá repozitář metadat, do kterých se ukládají informace o databázi. Posledním prvkem je dotazovací jazyk, kterým je možné přistupovat k datům. (4, s. 1-1)

3.2.1 Oracle Database

Oracle Database je databázový systém vytvořený společností Oracle. V roce 1979 byla představena Oracle Database V2, jako první komerční databázový systém využívající strukturovaný dotazovací jazyk (SQL). Společnost Oracle od té doby vydala 11 verzí jejich databáze a to od V2 až po 12c. Každá verze přinesla vylepšení nebo přídavky jako je například PL/SQL jazyk, změna z relační technologie na objektově-relační, souborový systém nebo také umožnění připojení více instancí k jedné databázi a tím rozložení zátěže. (4, s. 1-3)

Relační databázové systémy mají výhodu v nezávislosti fyzických a logických struktur. V Oracle Database jsou kolekce logických struktur, nebo objektů nazývány schémata. Tato schémata jsou navázána na uživatele a mají stejný název. Objekty těchto schémat jsou objekty vytvořené přímo uživateli a odkazují na data v databázi. Schémata jsou jedním z typů databázových objektů. (4, s. 1-4)

PL/SQL je procedurální nadstavbou nad SQL, umožňující používat všechny SQL dotazy, funkce a datové typy, ale navíc přidává možnost kontroly SQL programů, používání proměnných nebo správu chyb. Hlavní výhodou PL/SQL je schopnost ukládat aplikační logiku přímo do databáze. PL/SQL procedury nebo funkce patří do objektů schémat, které se skládají z SQL dotazů, spojených dohromady a spuštěny najednou pro splnění požadovaného cíle. (4, 1-5)

Strukturu Oracle Database je možné brát z fyzického tak i logického pohledu. Do fyzické struktury patří soubory a procesy, které se zobrazují na úrovni operačního systému. Oproti tomu logická struktura je významná pouze pro databázový systém. Tabulka nebo tabulkové prostory jsou objekty na úrovni databázového systému a je možné s nimi pracovat pouze na této úrovni. (4, s. 1-8)

Oracle Database využívá pro fyzickou strukturu datové soubory značené „**datafile**“, kontrolní soubory značené „**controlfile**“, pro transakční záznamy „**redolog**“ a pro archivaci transakcí „**archivelog**“. S ohledem na téma bakalářské práce, je nutné znát z pohledu logické struktury pouze tabulkový prostor „**tablespace**“. (4, 1-9)

Důležitou součástí Oracle Database je prostor rychlého obnovení „**flash recovery area**“. Jedná se o centrální repozitář dané databáze pro všechny soubory související se zálohováním a obnovou. Prostor rychlého obnovení, takzvaně FRA uchovává tyto typy souborů:

- Zálohové soubory
- Zálohy archivních záznamů
- Archivní záznamy
- Zálohy kontrolních souborů
- Používané soubory transakčních záznamů
- Používané kontrolní soubory

(3, s. 159)

3.2.2 Microsoft SQL Server

Microsoft SQL Server je relační databázový systém vyvinutý společností Microsoft. První verzí byl Microsoft SQL Server 1.0 vydaný roku 1989. Do roku 2000 bylo vydáno celkem 5 verzí tohoto databázového systému. Verze vydané po roce 2000 jsou značeny rokem jejich vydání. (2)

Logické struktury jsou nazývány schémata. Schémata jsou kolekce databázových entit, které společně formují jmenné prostory „**namespace**“. Schémata také mohou obsahovat tabulky, které jsou vytvořeny pod jinými schématy. Microsoft SQL Server používá pro fyzickou strukturu datové soubory nazývané „**datafile**“, pro transakční záznamy

„**transaction log**“ a pro archivaci transakcí „**transaction log backup**“. Jako logickou strukturu je používán termín „**filegroups**“. (2, s. 697)

3.3 Zálhování a archivace

Společnost Microsoft definuje zálhování jako proces vytvoření plnohodnotné kopie databáze, případně jejích transakčních záznamů na hardwarově nezávislé uložení za účelem vytvoření datové zálohy nebo záznamové zálohy. (5, s. 7)

Výsledkem procesu zálhování je zálhový soubor. Zálhový soubor je chápán jako datový soubor, ze kterého je možné obnovit databázi po poruše. Tento soubor je poté nutné uschovat na bezpečném a chráněném uložení, aby se předešlo jeho ztrátě, poškození, zneužití nebo nedovolené změně. Požadavky na uchování zálhových souborů budou popsány v pozdější kapitole. (5, s. 7)

Je nutné ovšem rozlišovat pojmy zálhování a archivace. Výsledek zálohy jsou aktivní kopie provozních dat, které jsou použity k případné obnově. Oproti tomu výsledkem archivace je kopie dat, které již nejsou provozní, ale jejich kopie musí být stále k dispozici. Jedním z hlavních rozdílů těchto pojmů je doba ukládání. Zálhové kopie jsou ukládány pouze po dobu zálhovacího cyklu a zpravidla jsou uloženy v blízkosti hlavního systému pro dosažení co nejrychlejšího obnovení v případě ztráty nebo poškození dat. Archivní kopie jsou uchovávány zpravidla na delší dobu, případně na vždy, a měly by být uloženy mimo objekt kde je umístěn hlavní systém. (6)

3.4 Rozdělení záloh

Zálohy je možné rozdělit na několik typů, které se rozdělují z několika kritérií. Typy záloh jsou popsány v této kapitole.

3.4.1 Komprimace

Dle tohoto kritéria je možné rozdělit zálohy z hlediska zpracování a to na komprimované a nekomprimované. Komprimace je proces zmenšování objemu dat pro

snížení požadavku na uložení pomocí předem zvoleného kompresního algoritmu. Při komprimaci je definován kompresní poměr, který určuje, o kolik se sníží výsledná velikost souboru. Poměr je možné vyjádřit procentuálně nebo v počtu bitů na byte. (7, s. 2)

Komprimace se rozlišuje na dva typy a to ztrátová a bezztrátová. Ztrátová komprimace označuje proces, při kterém jsou některá výchozí data pozměněna, případně nevýznamná data zanedbána, za účelem dosažení vyššího kompresního poměru. Dekomprimací se získají poněkud jiné hodnoty, než které byly komprimovány. Z toho důvodu se tento typ příliš nevyužívá v databázovém pojetí. Bezeztrátová metoda komprimace nemění ani nezanedbává jakékoliv data. Po dekomprimaci jsou získány totožná data s těmi, které byly původně komprimovány. (7, s. 2)

V průběhu zálohování databázových souborů je možné určit, zda se daná záloha má komprimovat, případně určit, že každá záloha bude komprimována. Také je možné zvolit, jaký kompresní algoritmus se má při záloze použít. (4, s. 171)

Využití komprimace záloh tedy umožňuje snížit nároky na prostor uložení, případně síťové prvky, pokud jsou zálohy přesouvány na jiné místo. Komprimací ovšem vzniká problém delšího trvání procesu zálohování, který se prodlužuje při každém zvýšení kompresního poměru. (7, s. 2)

3.4.2 Plná a částečná záloha

V tomto kritériu je možné rozdělit zálohy na dva typy, které se liší částí databáze, kterou zálohují. Jedná se o zálohu celé databáze a záloha částí databáze. Záloha celé databáze v sobě shromažďuje všechny datové soubory a alespoň jeden kontrolní soubor. Oproti tomu záloha částí databáze může obsahovat pouze specifikované datové soubory. Kontrolní soubory při částečné záloze nemusí být žádné. (8, s. 15-6)

3.4.3 Přírůstek

Kritérium přírůstku nerozlišuje typ záloh v úplném kontrastu. Zálohy dělené dle tohoto kritéria jsou spíše doplňující.

Přírůstkové kritérium udává dva typy záloh. Jedním z nich je plná záloha, kdy jsou zálohovány všechny datové bloky, které obsahují data. Druhým typem je takzvaná přírůstková záloha, která zálohuje datové bloky, které byly změněny od poslední zálohy.

Způsob provádění přírůstku se také rozlišuje. Jedná se o kumulativní a diferenciální zálohu. Kumulativní záloha zálohuje všechny změny od poslední plné zálohy. Diferenciální záloha oproti tomu zálohuje všechny změny pouze od poslední přírůstkové nebo plné zálohy. (8, s. 15-6)

3.4.4 Dostupnost databáze v průběhu zálohování

Kritérium dostupnosti databáze určuje v jakém stavu je databáze a zda jsou data v datových souborech konzistentní. Jedná se o online a offline zálohu. Při online záloze není záruka, že jsou data konzistentní. Důvodem vzniku tohoto problému je to, že data a kontrolní soubory nejsou synchronní. Z toho důvodu je nutné přepnout databázi do zálohovacího režimu, který přepne databázi do režimu zapisování dat do transakčních záznamů nikoliv přímo do datových souborů. Po ukončení zálohy datových souborů se následně zálohují transakční záznamy. Výhodou této zálohy je dostupnost databáze v průběhu zálohování. Při offline záloze se musí databáze uzavřít a tedy znemožnit přístup uživatelům k datům. Offline záloha má konzistentní data z důvodu synchronizace dat a kontrolních souborů. Nevýhodou tohoto typu je nedostupnost databáze v průběhu zálohování. (8, s. 15-6)

3.5 Požadavky na zálohové soubory

3.5.1 Ochrana

Ochranu záloh je možné rozdělit na dva druhy. Fyzické a logické zabezpečení přístupu k médiím, na kterých jsou uloženy soubory záloh. Předpokladem k ochraně dat jako takové je nepřístupnost těchto specifických dat velkému množství osob.

V případě fyzického zabezpečení se jedná o kontrolu přístupu již u vstupních dveřích do objektu. Dalším stupněm ochrany může být vrátný, který kontroluje, zda je osoba oprávněna ke vstupu. Jedním ze stupňů ochrany jsou také automatické dveře, které fungují na principu čipových nebo magnetických karet, případně numerického kódu. Dalšími prvky mohou být různé zařízení kontrolující narušení objektu, jako jsou snímače pohybu, snímače oken a dveří a kamerové systémy. Těmito způsoby je zabezpečena místnost, ve které jsou data, jak aktivní tak zálohy a archivy. V případě aktivních dat, je nutné zabezpečit celý počítačový systém. Tímto posledním stupněm ochrany jsou skříně. Obyčejná skříně je z hlediska ochrany nepraktická, vniknutí do vnitřku této skříně je velmi snadné. Náhradou jsou speciální bezpečnostní skříně například trezory, které mohou více odolávat útokům. (9, s. 52)

Důležitým prvkem je také fyzické oddělení aktivní databáze a zálohových souborů. V případě incidentu počítačového systému, na kterém je provozována databáze, je zajištěna ochrana zálohových souborů. V opačném případě by zálohové soubory nebylo možné získat, natož obnovit. (5, s. 6)

Problematiku logického zabezpečení přístupu k médiím ve většině případech řeší samotný operační systém. Ten se snaží hlídat přístup k datům z hlediska oprávnění uživatelů tak, aby k nim neměl přístup uživatel, který k nim nemá přístupová práva. Z toho důvodu je nejdříve nutné ověřit identitu uživatele a následně důkaz, že osoba vydávající se za tohoto uživatele je skutečný uživatel. K tomu slouží několik způsobů autentizace. Prvním je důkaz znalostí, kdy uživatel zadá heslo. Dalším způsobem je důkaz vlastnictví, kdy je důkazem povětšinou bezpečnostní karta nebo čip. Posledním způsobem je důkaz vlastnosti, kdy se uživatel prokáže svou unikátní tělesnou vlastností, například otiskem prstu. Tato vlastnost se také nazývá biometrika. Pomocí těchto důkazů může systém autentizovat uživatele s určitým přístupem k datům. Systém má také za úkol uživateli povolit pouze určitý druh operací. Může se jednat pouze o čtení, nebo čtení a zápis, případně smazání. (9, s. 56)

Další způsob logického zabezpečení je kryptografie, konkrétně šifrování. Způsobů je hned několik. Jedním z nich je offline šifrování souborů. Systém zašifruje pouze určené soubory. Vylepšením tohoto způsobu je online šifrování souborů. V tomto případě systém

šifruje všechny soubory, které splní určená kritéria. Kritériem může být například určitá složka nebo soubory odeslané elektronickou poštou. Poslední možností je šifrování celého disku. Systém šifruje veškerý tok dat směřující na uložení. Nevýhodou jakéhokoliv šifrování je obtížný způsob nakládání s hesly nebo šifrovacími klíči a vysoké využití procesoru. (9, s. 58)

Pro šifrování zálohových souborů i databáze je možné zvolit metodu, jakou se má zašifrování provést. Jednou z nich je základní způsob. Ten využívá různých ochranných metod, jako jsou hesla, klíče nebo certifikáty, které jsou nutné pro dešifrování. Další způsob využívá pouze hesla a je nutné ho specifikovat při spouštění zálohování.¹ (3, s. 172)

Chránit data je nutné také při přenosu dat, ve chvíli kdy data opustí chráněné prostředí. Jakmile data vstupují do vnější sítě, může je kdokoliv odchytnout a zkompromitovat nebo modifikovat. Proto je nutné využít způsobů šifrování, případně digitálních otisků. Digitální otisk je kontrolní součet, který je vypočten před odesláním dat. Po přijetí dat se znovu spočítá kontrolní součet, a pokud tyto součty souhlasí, data jsou nemodifikovaná. (9, s. 59)

3.5.2 Duplicita

Duplicita doplňuje možnosti ochrany záloh. Přesto se jedná o samostatný požadavek a vyjadřuje počet identických kopií jednotlivých záloh. Čím více kopií, tím více klesá pravděpodobnost ztráty dat. Všechny kopie zálohy ovšem musí být stejně zabezpečeny. Pokud je jedna z kopií chráněna méně, než zbylé kopie, tak se celková ochrana dat snižuje.

Problém duplicitních dat nastává ve vyšším požadavku na uložení a na jejich synchronizaci. Pokud je jedna z kopií zastaralá, stává se nepoužitelnou. Proto je nutné udržovat všechny kopie identické. (9, s. 63)

¹ Terminologie metod šifrování se liší dle společnosti, která vyvíjí daný databázový systém.

3.5.3 Fyzické rozdělení záloh dat a záznamů

Rozdělení záloh v geografickém smyslu je snadným způsobem, jak předejít ztrátě dat kdy jsou na vině přírodní katastrofy. Rozdělení záloh je také jedním ze způsobů ochrany, který je možné kombinovat s duplicitou. (9, s. 62)

3.6 Validace záloh

Validace je proces, při kterém se kontrolují zálohové soubory. Důvodem je zjištění, zda některý ze souborů nebyl fyzicky porušen a zda je možné jej použít k obnově dat. Tato kontrolu ovšem není nutná, záleží pouze na důslednosti databázového administrátora, zda bude validovat zálohové soubory před jejich využitím. Díky validaci je možné zjistit, že některé zálohové soubory nebude možné použít a včas se tomu přizpůsobit.

Další možností je validaci záznamů o zálohách. Při tomto procesu se kontrolují záznamy zálohovacího systému proti souborům na uložišti a zároveň je proveden kontrolní součet. Včasným použitím této metody je možné předejít problémům s obnovou. (3, s. 184)

3.7 Obnova

Obnova je proces, při kterém je databáze navrácena do stabilního a konzistentního stavu s využitím zálohových souborů vytvořených při procesu zálohování. Tento proces má tři fáze. První fází je přepnutí porušeného objektu do offline stavu, kdy se aktivní porušená data přepisují daty ze zálohy. Druhou fází je obnovení dat. Tato část, nazývaná obnova, zkopíruje všechny data a záznamy ze zálohových souborů do aktivní databázové evidence. Třetí fází je zotavení dat. V této fázi jsou data posunuty v čase pomocí transakčních záznamů pro zachování konzistence a dosažení stejného stavu v jakém byla původní data. (5, s. 7)

3.7.1 Politiky udržování záloh

Okno obnovení je způsob, kterým se specifikuje doba, po kterou chceme udržet zálohy na diskovém uložišti. Tento způsob je založen na stanovení nejdelší zpětné doby, do kdy má být možné obnovit databázi. Při stanovení doby, po které se vytváří zálohy, je nutné brát v potaz právě okno obnovení. Pokud, například, bude nastaveno okno obnovení na dva

týdny a vytváření plných záloh bude s týdenním intervalem, je nutné uchovat tři plné zálohy. (10, s. 405)

Druhým způsobem je takzvané redundantní udržení záloh. Tento způsob je založen na stanovení počtu minulých plných záloh, které se mají zachovat. Jakmile je vytvořena plná záloha, uchová se požadovaný počet záloh a zastaralé je možné smazat. (10, s. 406)

3.8 Strategie zálohování a obnovy

Strategie zálohování a obnovy je dokument s navrženým postupem zálohování a archivace dat vzhledem k požadavkům zákazníka. Tato strategie také obsahuje postupy pro obnovení po incidentu. (10, s. 6)

Část o zálohování a archivaci dat určuje typy a frekvenci vytváření záloh, hardware, který musí být schopen splnit tyto požadavky, způsob testování zálohových souborů a jak mají být tyto soubory ukládány. Část o obnově specifikuje frekvence testování obnovy pro zjištění obnovitelnosti ze zálohových souborů například na testovací databázi. V části o obnovování je také určena osoba zodpovědná za proces obnovy, způsob obnovy k dosažení cílů dostupnosti a minimalizování ztráty dat.

Vytváření tohoto dokumentu vyžaduje pečlivé plánování a testování strategie ve všech možných naplánovaných kombinacích. Testování je nutné pro potvrzení funkčnosti strategie. Při vytváření je nutné brát v potaz požadavky na databázovou evidenci, zejména požadavky na dostupnost a ochranu dat. Zvážit je nutné také povahu a vlastnosti každé databáze, jako jsou velikosti, vzorce používání, povahu obsahu a podobné. Také je nutné se zabývat omezením zdrojů, jako jsou lidé, výkonnost počítačových systémů, uložště pro zálohové soubory a jejich ochranu. (5, s. 8)

Při vytváření optimální strategie zálohování je nutné zpočátku zodpovědět specifické otázky. Tyto otázky jsou následující:

- Kolik hodin denně je pro aplikace nutné připojení k databázi?
 - o V případě, že je očekáván určitý časový interval, kdy není k databázi vysoký přístup, je výhodné naplánovat zálohu celé databáze právě do tohoto intervalu.

- Jak často se vkládají a mění data, případně vymazávají?
 - o Na základě tohoto intervalu se určí, jakým způsobem jsou prováděny přírůstkové zálohy.
- Vyskytují se změny v celé databázi nebo jen v určitých částech?
- Jak velké uložení bude potřeba pro celkovou zálohu?

Zodpovězením těchto otázek je možné předejít problém, které se objeví až v průběhu zálohování. Například příliš málo místa na uložení nebo nedostatečný výkon zpracování požadavků. (5, s. 9)

3.9 Zálohovací nástroje

3.9.1 Recovery Manager – Oracle Database

V prvních verzích Oracle Database bylo zálohování řešeno pomocí přímého exportu na magnetické pásky, nebo použitím Enterprise Backup Utility. Recovery Manager, zkráceně RMAN, byl poprvé použit s Oracle Database verze 8.0.3. Od té doby je součástí každého vydání Oracle Database. V každé další verzi jsou implementovány nové funkce, případně upraveny staré. (10, s. XXVI)

RMAN je samostatná aplikace, která se připojuje k databázi jako klient pro získání přístupu k interním zálohovacím a obnovovacím balíčků. Rozhraní je velmi jednoduché. Příkazy jsou vkládány do příkazové řádky, kde jsou interpretovány a odeslány do databázového systému. RMAN má na starost pouze koordinování událostí, samotné zpracování interpretovaných příkazů probíhá přímo v databázovém systému. (10, s. 34)

Pro správnou funkčnost nástroje RMAN je vždy potřeba specifikovat cílovou databázi „**target**“ kam budou posílány příkazy. Dodatečnou funkcionalitu poskytuje katalog. Pojem cílová databáze je pouze jiný název pro databázi, ke které je RMAN připojen a veškeré zálohování nebo obnova bude provedeno právě na této databázi. (10, s. 34)

Připojení RMAN k cílovému databázovému systému probíhá na nejvyšší úrovni, s oprávněním administrátora. Pro RMAN jsou nutné oprávnění jako zapnutí a vypnutí databáze, v případě obnovy i její vytvoření. (10, s. 35)

Velmi důležitou součástí RMAN je katalog. Jedná se o databázové schéma, kam jsou ukládány veškeré informace o RMAN zálohách. Na rozdíl od kontrolního souboru nejsou informace v katalogu mazány v pravidelných intervalech. Použití katalogu má několik výhod:

- Ukládání RMAN skriptů
- Uchování informací o zálohách napříč celým databázovým systémem.
- Vyšší flexibilita při vytváření zpráv o průběhu zálohování.

(10, s. 208)

3.9.2 Management Studio – Microsoft SQL Server

Management Studio ve verzi SQL Serveru 2005 nahradilo původní nástroj Enterprise Manager. Tento nástroj je velmi cenným prostředkem databázového administrátora ke správě databáze a sledování zatížení. Mezi činnosti, které tento nástroj dokáže provádět, patří konfigurace vlastností jádra databázového systému, vytváření a správa databáze, tabulek a jiných objektů, vytváření úloh, provádění záloh, reportování výkonu a operací serveru, monitorování aktivity a mnoho dalších činností. (2, s. 93)

3.10 Principy zálohování a obnovy

3.10.1 Export a import

Nástroje export a import jsou používány k uložení dat z databázových souborů do souborů operačního systému v binárním formátu. Tyto soubory obsahují tabulkové záznamy, vlastnosti objektů a kontrolní informace. Soubory exportované lze následně importovat jak do původní tak i jiné databáze. Při exportu a importu však nemůže být využit princip transakčních záznamů. (4, s. 18-7)

3.10.2 Manuální záloha a obnova

Tento typ zálohování je pro administrátora nejvíce zatěžující. Je nutné, aby dodržel určitou posloupnost příkazů, díky kterým se databáze zpřístupní pro zálohování nebo obnovu. Samotné zálohování a obnova jsou prováděny pomocí příkazů operačního systému, zejména kopírování. Databáze se na pokyn administrátora přepne do zálohovacího módu a zanechá datové soubory v konzistentním stavu. Během této doby se všechny operace provádí pouze v paměti. Jakmile administrátor ukončí zálohovací mód, do originálních datových souborů se převedou všechny operace v paměti. Administrátor si také musí pamatovat veškeré informace o zálohových souborech. (11, s. 28-1.)

3.10.3 Specializovaný nástroj

Tyto nástroje přináší administrátorům velmi jednoduchý způsob zálohování a obnovy. Vedou si záznamy o zálohované databázi, ukládají a validují zálohové soubory. Způsob ukládání všech těchto informací se liší v každém databázovém systému. V Oracle Database s pomocí nástroje RMAN jsou tyto způsoby dva. Jedním je uložení do katalogu a tím druhým je ukládání do kontrolního souboru. Tyto nástroje samostatně obsluhují přepínání databáze, ukládání souborů a jejich vlastností do svých záznamů a také jejich mazání. (11, s. 2-1)

4 Praktická část

Veškeré nastavení zálohy a obnovy, jejich spouštění a testování časové náročnosti bude prováděno na testovacím serveru s Oracle Database 11gR2. Jedním z faktorů při výběru tohoto databázového systému byla autorova zkušenost. Další vliv na výběr měli i jiné výhody. Mezi ně lze zařadit vysokou rozšířenost ve společnostech, možnost instalace a provozu na většině dostupných operačních systémech a velmi slušné uplatnění administrátorů Oracle Database na pracovním trhu.

4.1 Vliv datového nárůstu na velikost záloh

Před navržení strategie zálohování je výhodné znát co nejvíce informací o průběhu zálohování. Tyto informace se týkají časové náročnosti a požadavku na uložení. Získání těchto informací bude provedeno na testovací databázi. Naplnění této databáze bude provedeno pomocí skriptů, jejichž spuštění bude naplánováno pomocí příkazu „crontab –e“ a následnému načasování spuštění. Tím je možné nasimulovat přístupy a vkládání dat v reálné společnosti. Samotné zálohování je realizováno pomocí specializovaného nástroje RMAN.

Naplánování spuštění zálohování pro zjištění většího množství výsledků testování:

- Plná záloha je prováděna 1x denně o půlnoci.
- Přírůstkové zálohy jsou naplánovány na každou lichou hodinu.
- Záloha archivních logů je spuštěna každých 15 minut, každou hodinu.
- Skript na plnění databáze je proveden každých 10 minut, každou hodinu.
- Testovací tabulkový prostor je nastaven na 1024 MB.
- Všechny zálohy jsou nekomprimované.

Analýza náročnosti byla provedena na testovacím serveru společnosti Your System s.r.o s touto konfigurací:

- Čtyř jádrový procesor Intel Xeon E5430 s frekvencí 2,66 GHz
- Operační paměť 3,5 GB a odkládací prostor 3 GB
- Uložení se skládá ze čtyř disků. Jeden 100 GB disk, dva 32 GB a jeden 10 GB.

4.1.1 Plná záloha

První nekomprimované zálohování celé databáze vytvořilo soubor o velikosti 1,02 GB. V testovacím tabulkovém prostoru bylo v průběhu testování zaplněno 425 MB prostoru. Plné zálohy se postupně zvětšovali společně s využitím tabulkového prostoru. Z této úměrnosti tedy vyplývá, že potřebný prostor pro ukládání záloh musí být dostatečně velký s ohledem na požadovaný počet redundantních kopií, případně na požadované okno obnovení.

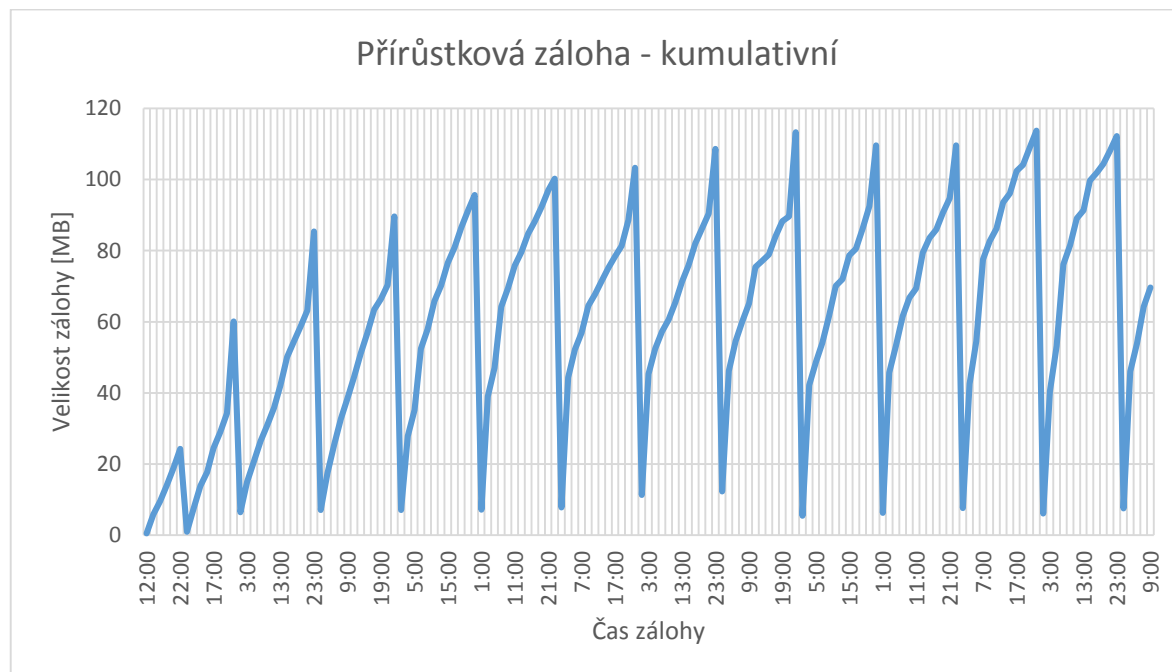


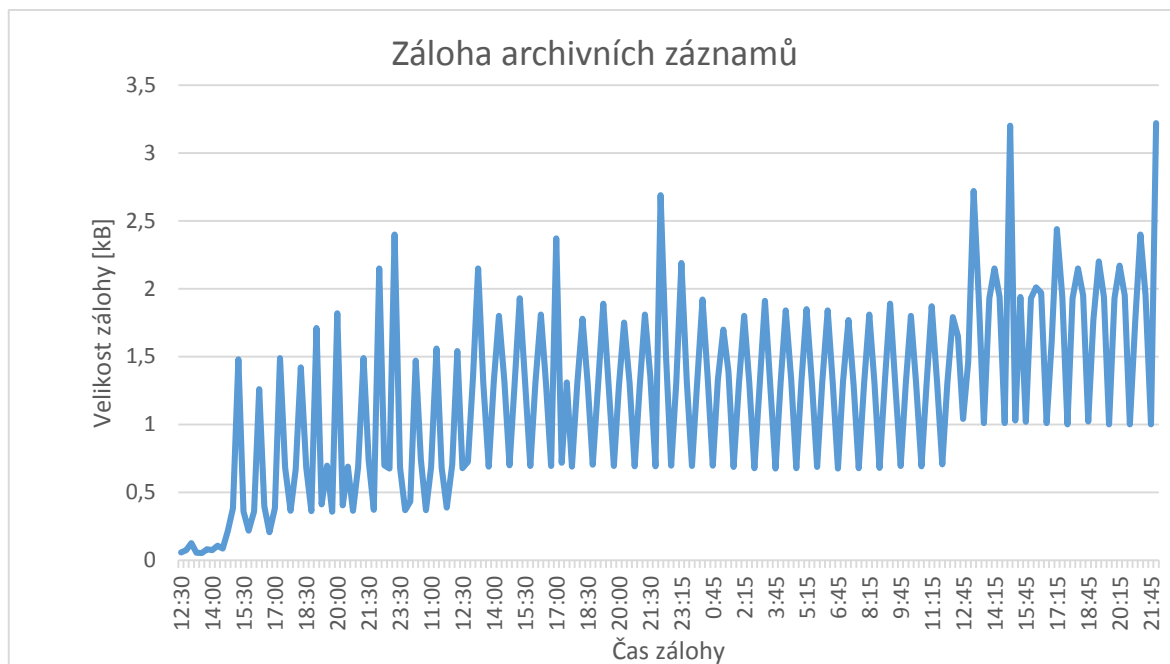
Obrázek č. 1 - Vývoj velikosti souborů plné zálohy.
Zdroj: Autor

Obrázek č. 1 zobrazuje nárůst velikosti plné zálohy v průběhu měření. Mezi dny 13. 9. a 14.9 nebyla databáze zatěžována z důvodu příliš malého plnění daty do pouze jedné tabulky. Další den byly přidány dvě další tabulky a skript na vkládání do tabulek byl rozšířen. Z původních 1000 záznamů každých 10 minut, bylo vkládání zvýšeno na 3000 záznamů. Po dobu testování bylo do původní tabulky vložena přibližně 17 milionů záznamů a do zbylých dvou po 15 milionech záznamů. Celkem bylo tedy vloženo téměř 50 milionů záznamů. V tabulkách jsou dva textové atributy o 30 znacích a jeden číselný atribut o 9 znacích.

4.1.2 Přírůstkové kumulativní zálohy

Přírůstkové zálohy byly prováděny každou lichou hodinu. V průběhu testování bylo vytvořeno 150 zálohových souborů.





Obrázek č. 3 – Vývoj velikosti souborů zálohy archivních záznamů

Zdroj: Autor

Obrázek č. 3 zobrazuje pouze 188 údajů z důvodu nepřesažení pozorovatelného množství. S vyšším počtem se stává obrázek značně nepřehledný. Z průběhu grafu lze identifikovat časy, kdy byly přidány tabulky a znásobeny skripty na plnění těchto tabulek.

4.2 Analýza současné situace

V této podkapitole je popsána charakteristika společností, zálohovací strategie a následné porovnání a zhodnocení těchto strategií.

4.2.1 Charakteristika společností

Společnost A je velká společnost, která provozuje přibližně 80 databází na databázovém systému Oracle verze 11gR2 Enterprise Edition. Do těchto databází přistupují dva typy aplikací. Interní aplikace pro fungování společnosti, kterou využívá téměř 1800 zaměstnanců a webová aplikace pro téměř 500 tisíc uživatelů. Také zde fungují desítky jiných aplikací. Pro zálohování je používán nástroj RMAN.

Společnost B je menší společnost provozující 10 databází na databázovém systému Oracle verze 11gR2 Enterprise Edition. Interní aplikaci používá 250 zaměstnanců. Dalším

způsobem využívání databází je datové centrum webového portálu. Pro zálohování je používán nástroj RMAN.

4.2.2 Strategie zálohování

Společnost A má nastaveno zálohování archivních záznamů jednou denně. Přírůstkové zálohy se neprovádí a záloha celé databáze je prováděna jednou týdně.

Zálohy jsou ukládány na sériovou pásku a na diskové uložení. Při zálohování archivních záznamů jsou všechny archivní soubory smazány. Je nakonfigurována automatická záloha kontrolního souboru.

Ve společnosti B jsou archivní záznamy zálohovány každých 15 minut. Vytváření kumulativních přírůstkových záloh je prováděno každý den, mimo sobotu. Sobota je vyhrazena pro zálohování celé databáze a vytvoření kopie této zálohy. Zálohy jsou ukládány na diskové uložení.

4.2.3 Zhodnocení strategií zálohování

Tabulka č. 1 – Denní frekvence zálohování

Společnost	Archivní záznamy	Přírůstkové zálohy	Plná záloha	Kopie
A	1x	0x	1x	1
B	96x	1x	1x	2

Zdroj: Autor

Ze získaných informací je možné zpozorovat ohromný rozdíl ve frekvenci zálohování archivních záznamů jednotlivých společností. V případě nechtěné události je pro společnost A velmi pravděpodobná ztráta dat z celého dne. Ve společnosti B je tato pravděpodobnost omezena na pouhých 15 minut. Společnost A podstupuje ohromné riziko a je možné, že v případě havárie budou ztracena data od poslední zálohy archivních záznamů, tedy data z celého dne.

Společnost B má podstatně lepší zabezpečení svých dat, co se týče zálohování archivních záznamů a přírůstkových záloh. Společnost B také lépe obstarává duplikaci záloh. Z logického pohledu je tedy výrazně lépe zhotovena strategie zálohování ve společnosti B.

4.3 Identifikace požadavků

Při analýze současné situace bylo zjištěno několik kritérií, dle kterých je možné identifikovat požadavky na zálohování. Tyto požadavky jsou následující:

- Frekvence záloh archivních záznamů
- Frekvence přírůstkových záloh
- Frekvence plných záloh
- Umístění jednotlivých záloh
- Počet kopií jednotlivých záloh

Další požadavky mohou být identifikovány na základě teoretické části. Tyto požadavky byly již specifikovány v kapitole 3.5.

4.4 Definice společnosti

Pro co nejlepší navržení strategie zálohování je nutné vědět pro jakou společnost je strategie navrhována. Z toho důvodu je důležité znát charakteristiku společnosti, odvětví, ve kterém podniká, využití databázových evidencí ve společnosti a požadavky, na základě kterých bude strategie zálohování navržena.

4.4.1 Charakteristika společnosti

Fiktivní společnost, pro kterou je navrhována strategie, je společností malé velikosti. Ve společnosti pracuje přibližně 50 zaměstnanců na pozicích, které vyžadují využívání databázových evidencí. Databázová evidence běží na vlastním serveru společnosti. Byl vybrán databázový systém Oracle Database, kdy budou zároveň spuštěny 2 instance. Jedna z instancí je určena pro data, druhá pro vytvoření katalogu.

4.4.2 Požadavky na strategii zálohování

Společnost vyžaduje co nejčastější zálohování celé databáze vzhledem k počtu zaměstnanců a přístupům, které zároveň nebude příliš zatěžující pro databázový server. Dále je požadováno přírůstkové zálohování alespoň jednou denně. Pro zálohování archivních

záznamů je společnost rozhodnuta pro krátký interval. Celková velikost produkčních dat se odhaduje v řádu desítek gigabytů.

Archivaci záloh požaduje společnost na externí diskové uložení v době jednoho roku. Také je požadováno co nejrychlejší obnovení v případě havárie. Na základě odhadované velikosti dat je doba obnovení stanovena na maximálně 3 hodiny.

4.5 Návrh strategie zálohování

Z požadavků společnosti na zálohování je nutné sestavit strategii zálohování. Frekvence plných záloh bude stanovena na jednu zálohu týdně.

Přírůstkové zálohy budou vytvářeny jednou denně, vždy v 10 hodin večer, mimo dny kdy se vytváří plná záloha. Kopie budou vytvářeny na stejné diskové uložení jako kopie plných záloh.

Archivní záznamy budou zálohovány každou půl hodinu od 7 ráno do 7 večer. Mimo tento rozsah budou zálohovány každou hodinu. Z důvodu malé velikosti těchto záloh budou vytvářeny 2 kopie. Jedna na stejné diskové uložení jako kopie plných a přírůstkových záloh a druhá bude ukládána na další diskové uložení serveru.

Archivace záloh se v této bakalářské práci bude řešit pouze minimálně a krátkodobě. Je to z důvodu jednoduchosti a pouze malé části zálohovacího procesu.

Pro otestování časové náročnosti budou použity principy zálohování a obnovy. Zálohování pomocí specializovaného nástroje již bylo provedeno v kapitole 4.1, ovšem na jiném testovacím serveru a hlavním cílem bylo zjištění vlivu datového nárůstu na velikost přírůstkových záloh a záloh archivních záznamů. Z toho důvodu bude analýza náročnosti specializovaného nástroje provedena pouze na plné zálohy databáze. Také bude provedeno testování komprimace. Následně bude vyhodnocen nejvýhodnější princip a ten bude použit pro realizaci strategie zálohování.

4.6 Analýza náročnosti

Konfigurace testovacího serveru je následující:

- Čtyř jádrový procesor Intel i5-750 s frekvencí 2,67 GHz
- Operační paměť 4 GB a odkládací prostor 4 GB
- Uložiště se skládá ze dvou disků. Jeden 45 GB (Disk A) a druhý 55 GB (Disk B).

Na testovacím serveru je provozována Oracle Database 11g Enterprise Edition. Testovací databáze je BAP se stejnojmennou instancí. Databáze pro katalog je CATDB. Disk A je rozdělen na oddíly kořenového adresáře a adresáře /u01 až /u03 pro Oracle Database. Disk B s 55 GB je rozdělen na tři oddíly, 20 GB, 17 GB a 16 GB. Tyto oddíly jsou připojeny do adresáře /u04 a jsou používány pro testovací data a zálohy. Testování časové náročnosti bude provedeno při naplnění databáze daty o velikosti 500 MB, 1 GB, 5 GB a 10 GB.

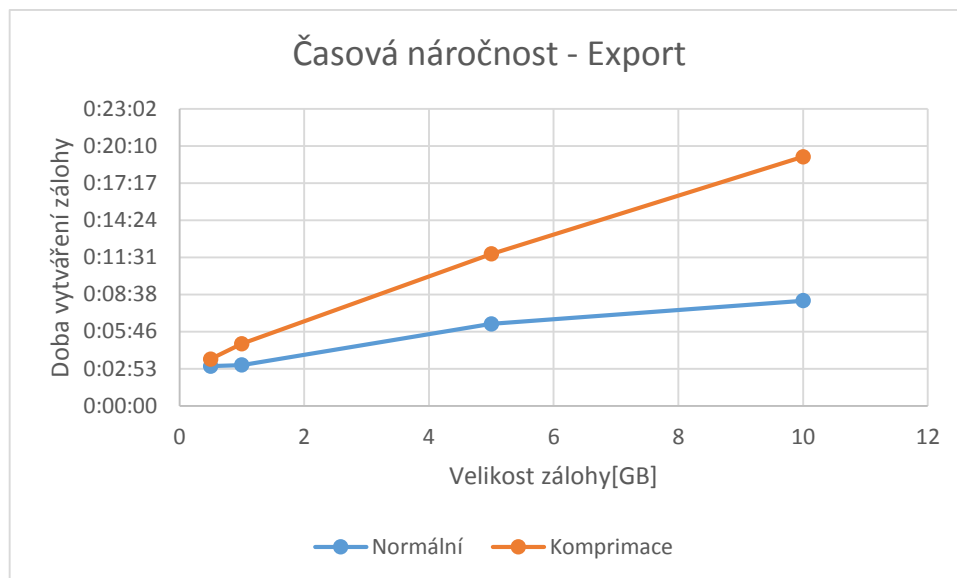
4.6.1 Export

Zálohování pomocí exportu probíhá použitím datové pumpy. Konkrétně se jedná o použití exportní funkce datové pumpy příkazem „expdp“. Před použitím datové pumpy je nejprve nutné specifikovat v databázi adresář, kam se budou data exportovat, případně se z tohoto adresáře data importují.

Tabulka č. 2 – Časová náročnost zálohování pomocí exportu

Databáze	Velikost	Čas	Velikost – komprim.	Čas – komprim.
500 MB	514 MB	00:03:05	345 MB	00:03:38
1 GB	930 MB	00:03:10	635 MB	00:04:49
5 GB	4,5 GB	00:06:22	3,1 GB	00:11:47
10 GB	8,8 GB	00:08:10	6,1 GB	00:19:19

Zdroj: Autor



Obrázek č. 4 – Graf časové náročnosti exportu
Zdroj: Autor

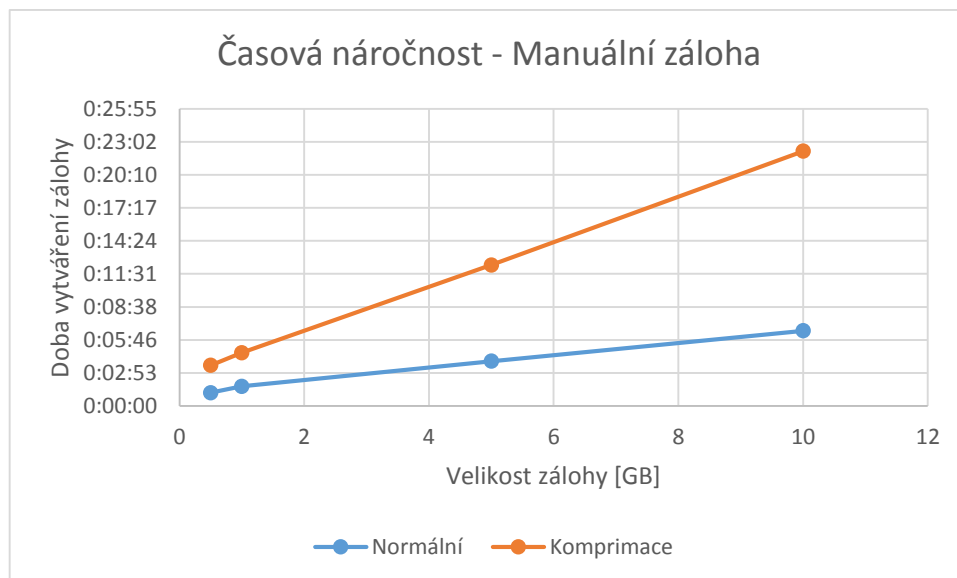
4.6.2 Manuální záloha

Vytvoření zálohy manuálně je prováděno jednoduchým kopírováním datových souborů. Manuální zálohování je prováděno po spuštění zálohovacího módu v databázi pomocí příkazu „ALTER TABLESPACE *tablespace* BEGIN BACKUP“, příkazu „ALTER DATABASE BEGIN BACKUP“ nebo po vypnutí celé databáze. Následně se kopírují datové soubory databáze a případně i komprimují. Po zkopírování datových souborů se ukončí zálohovací mód pomocí příkazu „ALTER TABLESPACE *tablespace* END BACKUP“ nebo „ALTER DATABASE END BACKUP“. V případě vypnutí se databáze znovu spustí. Archivní záznamy je také možné zálohovat pomocí prostého kopírování.

Tabulka č. 3 – Časová náročnost manuálního zálohování

Databáze	Velikost	Čas	Velikost – komprim.	Čas – komprim.
500 MB	2,8 GB	00:01:09	750 MB	00:03:32
1 GB	3,3 GB	00:01:43	1,1 GB	00:04:39
5 GB	7,2 GB	00:03:54	3,8 GB	00:12:18
10 GB	13,1 GB	00:06:33	7,3 GB	00:22:13

Zdroj: Autor



Obrázek č. 5 – Graf časové náročnosti manuální zálohy
Zdroj: Autor

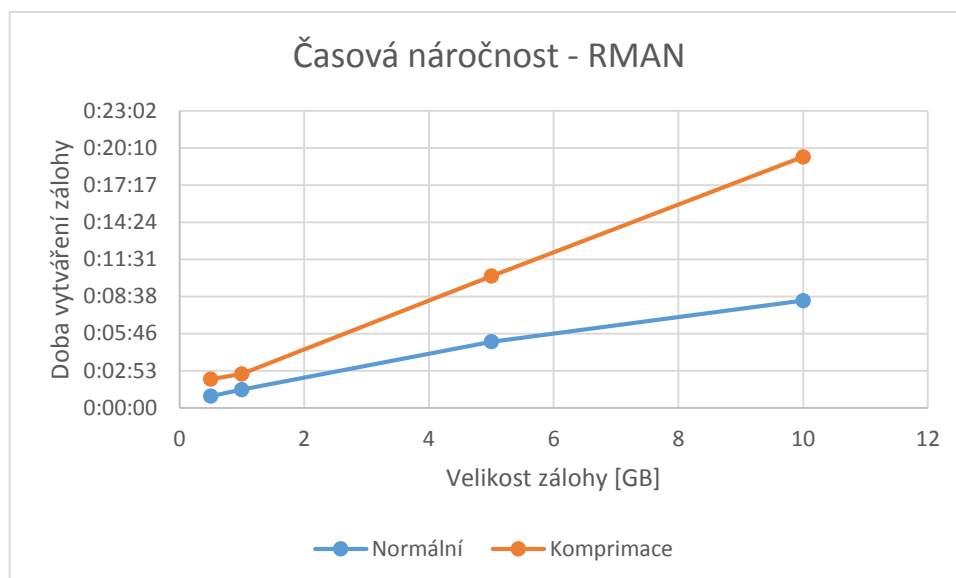
4.6.3 Specializovaný nástroj

Zálohování je provedeno nástrojem RMAN a příkazem „BACKUP AS BACKUPSET DATABASE“. Klauzule tohoto příkazu jsou například „COMPRESSED BACKUPSET“ pro zajištění komprimace zálohy.

Tabulka č. 4 – Časová náročnost zálohování pomocí RMAN

Databáze	Velikost	Čas	Velikost – komprim.	Čas – komprim.
500 MB	1,46 GB	00:00:55	558 MB	00:02:14
1 GB	1,96 GB	00:01:25	889 MB	00:02:38
5 GB	6,14 GB	00:05:08	3,40 GB	00:10:13
10 GB	10,6 GB	00:08:19	6,58 GB	00:19:28

Zdroj: Autor



Obrázek č. 6 – Graf časové náročnosti RMAN
Zdroj: Autor

4.7 Výběr optimálního principu zálohování

Na základě získaných poznatků z předchozí kapitoly 4.6 je nutné určit optimální postup zálohování. Výběr je založen na vícekritériální analýze variant. Kritériím, dle kterých se hodnotí, jsou přiděleny určité váhy. Váhy se mohou určit pomocí Saatyho matice. Výpočtem užítku je zjištěno, jaká varianta je nejvíce přínosná. Poté je daná varianta aplikována na strategii zálohování. Jako kritéria byla vybrána časová náročnost, náročnost na uložení, obnovitelnost co nejdříve k době nehody nebo výpadku a dostupnost databáze v průběhu zálohování. Do kritéria obnovitelnosti spadá několik vlastností, které budou ohodnoceny body. Tyto vlastnosti jsou validace záloh a obnovení archivních záznamů. Kritérium dostupnosti je také ohodnoceno body. V dostupnosti se hodnotí, zda je nutné databázi uvést do offline stavu, případně pouze požadované části databáze.

Kritéria a jejich váhy se liší dle odvětví, ve kterém je databáze nasazena. Je vhodné, pro tuto bakalářskou práci zavést pojmy „kritické systémy“ a „běžné systémy“. Databáze jako podpora kritických systémů jsou využívány v odvětvích, kde je kladen nejvyšší důraz na schopnost obnovit data co nejdříve k době nehody. Tato odvětví jsou například bankovníctví, letecký provoz nebo elektrárenský průmysl. Databáze jako podpora běžných systémů jsou používány v odvětvích, kde je dán větší důraz na náročnost na uložení. Do

těchto odvětví spadá většina běžných podniků, státní správa a podobné. Tyto odvětví spolu v některých případech mohou souviset a navzájem se doplňovat.

Saatyho matice pracuje s kvantitativním párovým porovnáním. Jedná se o expertní porovnání kritérií dle slovních definic. Pro kritické systémy je možné specifikovat tyto definice:

- Časová náročnost je slabě preferovaná před náročností na uložení
- Obnovitelnost je absolutně preferovaná před časovou náročností
- Obnovitelnost je absolutně preferovaná před náročností na uložení
- Obnovitelnost je silně preferovaná před dostupností
- Dostupnost je slabě preferovaná před časovou náročností
- Dostupnost je slabě preferovaná před náročností na uložení

Tabulka č. 5 – Saatyho matice pro odvětví kritických systémů

Kritické systémy	Čas	Velikost	Obnovitelnost	Dostupnost	Váha
Čas	1	3	1/9	1/3	22,56%
Velikost	1/3	1	1/9	1/3	17,59%
Obnovitelnost	9	9	1	5	34,39%
Dostupnost	3	3	1/5	1	25,45%

Zdroj: Autor

Pro běžné systémy je možné specifikovat tyto definice expertního porovnání:

- Časová náročnost je slabě preferována před náročností na uložení
- Obnovitelnost je slabě preferována před časovou náročností
- Obnovitelnost je silně preferována před náročností na uložení
- Dostupnost je silně preferována před časovou náročností
- Dostupnost je velmi silně preferována před náročností na uložení
- Obnovitelnost a dostupnost jsou rovnocenné

Tabulka č. 6 – Saatyho matice pro odvětví běžných systémů

Běžné systémy	Čas	Velikost	Obnovitelnost	Dostupnost	Váha
Čas	1	3	1/3	1/5	22,86%
Velikost	1/3	1	1/7	1/7	17,67%
Obnovitelnost	3	7	1	1	29,16%
Dostupnost	5	7	1	1	30,30%

Zdroj: Autor

Ze získaných vah se následně vypočítá užitek, který přináší jednotlivé principy zálohování. Výpočet je realizován skalárním součinem váhy a hodnoty daného kritéria principu.

Pro samotný výpočet užitku jsou hodnoty získány vzorcem: $\frac{y_i - D_i}{H_i - D_i}$, kde y_i je hodnota varianty daného kritéria, D_i je nejnižší hodnota daného kritéria a H_i je nejvyšší hodnota daného kritéria.

Tabulka č. 7 – Výpočet užitku pro kritické systémy

	Čas	Velikost	Obnovitelnost	Dostupnost	Celkem
Váha	22,56	17,59	34,39	25,45	
Export	0,08	1	0	0	0,19511
Manuální záloha	1	0	1	0	0,56952
RMAN	0	0,52	1	1	0,689865

Zdroj: Autor

Tabulka č. 8 – Výpočet užitku pro běžné systémy

	Čas	Velikost	Obnovitelnost	Dostupnost	Celkem
Váha	22,86	17,67	29,16	30,30	
Export	0,08	1	0	0	0,19614
Manuální záloha	1	0	1	0	0,520213
RMAN	0	0,52	1	1	0,686487

Zdroj: Autor

Tabulka č. 9 – Výsledné zhodnocení principů zálohování

	Kritické systémy	Běžné systémy	Celkem	Pořadí
Export	0,19511	0,19614	0,3912501	3
Manuální záloha	0,56952	0,520213	1,0897328	2
RMAN	0,689865	0,686487	1,376352	1

Zdroj: Autor

Optimální princip zálohování je identifikován a je možné nastavit samotné zálohování definované v zálohovací strategii z kapitoly 4.4.2.

4.8 Nastavení zálohování dle strategie

Zálohování bude prováděno nástrojem RMAN. Příkazy na vytváření záloh budou vloženy do skriptů. Časování spuštění a volání těchto skriptů bude nastaveno v cronu, plánovači úloh v unixových systémech. V této kapitole budou zobrazeny skripty a také výpisy cronu.

Následuje výpis skriptu „inc0_back.sh“, který po spuštění nastaví nutné proměnné, zkontroluje dostupnost adresáře, případně jej vytvoří a poté vytvoří plnou zálohu.

```
#Deklarace proměnných
export ORACLE_HOME=/u01/app/oracle/product/11.2.0/db_1
export ORACLE_BASE=/u01/app/oracle
export ORACLE_SID=BAP
echo "#####"
#Vytvoření adresáře pokud neexistuje
datum=$(date +%Y_%m_%d)
dir=/u04/backup/BAP/backupset/$datum
if [ ! -d "$dir" -a ! -h "$dir" ]; then
    echo "Creating directory"
    mkdir $dir
else
    echo "Directory exists"
fi
echo "#####"
#Vytvoření zálohy
$ORACLE_HOME/bin/rman target / catalog rman/"heslo"@catdb <<EOF
BACKUP AS BACKUPSET INCREMENTAL LEVEL 0 CUMULATIVE DATABASE
FORMAT
'/u04/backup/BAP/backupset/%Y_%M_%D/HOT0_%d_%t_%s_%c_%p.bkp';
exit
EOF
```

Samotná záloha je prováděna pomocí „INCREMENTAL LEVEL 0“ z důvodu kompatibility s „INCREMENTAL LEVEL 1“. Bez této klauzule není možné aplikovat princip přírůstkových záloh.

Následuje výpis skriptu „inc1_back.sh“. Tento skript nastaví proměnné a zkontroluje adresář a vytvoří přírůstkovou kumulativní zálohu.

```
#Deklarace proměnných
export ORACLE_HOME=/u01/app/oracle/product/11.2.0/db_1
export ORACLE_BASE=/u01/app/oracle
export ORACLE_SID=BAP
echo "#####"
#Vytvoření adresáře pokud neexistuje
datum=$(date +%Y_%m_%d)
dir=/u04/backup/BAP/backupset/$datum
if [ ! -d "$dir" -a ! -h "$dir" ]; then
    echo "Creating directory"
    mkdir $dir
else
    echo "Directory exists"
fi
echo "#####"
#Vytvoření zálohy
$ORACLE_HOME/bin/rman target / catalog rman/"heslo"@catdb <<EOF
BACKUP AS BACKUPSET INCREMENTAL LEVEL 1 CUMULATIVE DATABASE
FORMAT
'/u04/backup/BAP/backupset/%Y_%M_%D/HOT1_%d_%t_%s_%c_%p.bkp';
exit
EOF
```

Následuje výpis skriptu „arc_back.sh“. Na začátku jsou nastaveny proměnné a je zkontrolován adresář. Poté je přepnut soubor transakčního záznamu a soubory archivních záznamů jsou zálohovány.

```
#Deklarace proměnných
export ORACLE_HOME=/u01/app/oracle/product/11.2.0/db_1
export ORACLE_BASE=/u01/app/oracle
export ORACLE_SID=BAP
echo "#####"
#Vytvoření adresáře pokud neexistuje
datum=$(date +%Y_%m_%d)
dir=/u04/backup/BAP/arc_back/"$datum"
if [ ! -d "$dir" -a ! -h "$dir" ]; then
    echo "Creating directory"
    mkdir $dir
else
    echo "Directory exists"
fi
echo "#####"
#Přepnutí redolog
$ORACLE_HOME/bin/sqlplus -s / as sysdba <<EOF
alter system switch logfile;
```

```
EOF
#Vytvoření zálohy a smazání archivních záznamů
$ORACLE_HOME/bin/rman target / catalog rman/"heslo"@catdb <<EOF
  BACKUP ARCHIVELOG ALL DELETE INPUT FORMAT
  '/u04/backup/BAP/arc_back/%Y_%M_%D/arc_%t_%s_%h.bkp';
  exit
EOF
```

Následuje výpis časování z „cronu“. Každý spuštěný skript připiše do záznamového souboru.

```
#Plná záloha 18:30, sobota
30 18 * * 6 /home/oracle/BAPR/skripty/inc0_back.sh >>
/home/oracle/BAPR/log/inc0_back.log

#Přírůstková záloha 22:20, každý den kromě soboty
20 22 * * 0-5 /home/oracle/BAPR/skripty/incl_back.sh >>
/home/oracle/BAPR/log/incl_back.log

#Záloha arch. záznamů 7:05, 7:35, ... , 19:05, 19:35, každý den
5 35 7-19 * * * /home/oracle/BAPR/skripty/arc_back.sh >>
/home/oracle/BAPR/log/arc_back.log

#Záloha arch. záznamů 0:35, 1:35, ... , 6:35 a 20:35, ... , 23:35,
každý den
35, 0-6,20-23 * * * /home/oracle/BAPR/skripty/arc_back.sh >>
/home/oracle/BAPR/log/arc_back.log
```

Záznamy z provedení skriptů jsou zobrazeny v přílohách.

4.9 Zhodnocení řešení

Z výpisů specializovaného nástroje RMAN (Příloha č. 1 - Příloha č. 3) lze analyzovat vytvořené zálohové soubory. Soubor plné zálohy přibližně odpovídá výsledkům časové náročnosti získaných při testování a analýze náročnosti. Pro otestování zálohování archivních záznamů byla databáze naplněna daty o velikosti přibližně 120 MB. Údaj byl získán z dotazu pro výpočet použité velikosti prostoru rychlého obnovení. Vytvořený zálohový soubor tedy odpovídá archivním záznamům. Validace zálohových souborů, jak lze vypořádat z výpisu RMAN (Příloha č. 4 - Příloha č. 6), proběhla bez problémů.

Testovací obnova databáze proběhla taktéž v pořádku, jak je možné zjistit z výpisů (Příloha č. 7 - Příloha č. 8). Všechna data byla správně obnovena do původního stavu a to pomocí zotavení databáze k určenému pořadovému číslu.

4.10 Návrh vylepšení

Porovnáním strategie zálohování a obnovy a požadavků společnosti je možné identifikovat části zálohování, které je možné a vhodné vylepšit. V tomto případě se může jednat o vylepšení zálohovacích skriptů o vytvoření kopií zálohových souborů na definované externí uložení, kterým může být sériová páska. Pásky je poté možné fyzicky přenést do lépe chráněného prostředí. Dalším vylepšením může být kontrola, zda právě není spuštěn jiný zálohovací proces. Také by bylo vhodné zvýšit frekvenci zálohování archivních záznamů v době od 7 hodin ráno do 7 hodin večer. Určitým zpřehledněním by bylo rozdělení záznamů o průběhu skriptů pro zálohování archivních záznamů. Skript na zálohu archivních skriptů lze taktéž zjednodušit nastavením automatického přepínání souborů transakčních záznamů přímo v databázi po určeném časovém intervalu.

Jako další vylepšení by bylo možné přesunout databázi pro katalog na jiný server, aby v případě nehody provozního serveru byly všechny zálohové soubory stále přístupné. Je možné, že nastane nehoda na uložení, na kterém jsou datové soubory a zároveň zálohové soubory a poté nebude možné obnovit cílovou databázi.

5 Výsledky a diskuse

5.1 Zobecnění řešení

Řešení lze zobecnit pro další možná použití. V prvním kroku je vždy nutné identifikovat požadavky společnosti s ohledem na zálohování databázových evidencí. V dalším kroku je navržena strategie zálohování a obnovy na základě identifikovaných požadavků společnosti a se zřetelem na obecné požadavky na zálohy. V následujícím kroku je tato strategie otestována a v případě nedostatků je upravena. V posledním kroku je strategie implementována na požadované provozní, případně testovací databáze.

5.2 Diskuze

Navržené řešení může být ohroženo několika faktory. Malé zkušenosti databázového administrátora jsou jedním z nejvíce ohrožujících faktorů a to z důvodu, že nezkušený administrátor může znásobit vliv některých faktorů, které ohrozí navržené řešení. Mezi tyto další faktory patří nedostatečná frekvence zálohování archivních záznamů, nesprávná frekvence plného zálohování při použití okna obnovy. Do ohrožujících faktorů lze také započítat i malé diskové uložení určené pro zálohy, malou velikost operační paměti a například možnou vyšší poruchovost hardwaru, případně také výpadek propojení mezi databází a zálohovým uložištěm.

Předpoklady pro správnou funkcionalitu navrženého řešení jsou definované požadavky společností, otestování navržené strategie zálohování, konfigurace provozní, případně testovací databáze, konfigurace uložení určeného pro zálohové soubory a přístup do tohoto uložení.

Předpoklady také do jisté míry vychází z ohrožujících faktorů. Jedním z těchto předpokladů je zkušený administrátor, který zvládne navrženou strategii implementovat a v případě potřeby dokáže z vytvořených zálohových souborů spravovanou databázi obnovit.

Alternativním řešením může být zrcadlení databáze. V tomto režimu se transakční záznamy z primární databáze také provádí v „stand-by“ databázi. Tím je dosaženo uložení kopie dat. Při výpadku primární databáze se veškerý provoz přesune na stand-by databázi.

Avšak i v tomto režimu je nutné zálohovat jednu z databází, nebo obě pomocí principů definovaných v této bakalářské práci. Samotné zálohování je možné provádět pouze na stand-by databázi, díky čemuž nebude primární databáze zatěžována vytvářením zálohových souborů.

6 Závěr

Bakalářská práce měla za cíl navrhnout optimální strategii zálohování a obnovy, definovat požadavky kladené na tuto strategii a popsat teoretická východiska s ohledem na problematiku zálohování relačních databázových evidencí. V teoretické části byly popsány soubory databázového systému, objasnění termínů zálohování a archivace, dělení záloh, identifikované požadavky na zálohové soubory a validace záloh. Následně byl popsán proces obnovy, politiky udržování záloh, postup návrhu strategie zálohování a obnovy, specializované zálohovací nástroje databázových systémů a také samotné principy zálohování a obnovy.

V praktické části byla zmapována a zhodnocena současná situace zálohování ve dvou společnostech. Následně byla představena fiktivní společnost a její požadavky na zálohování a obnovu. Na základě těchto požadavků byla navržena strategie zálohování a obnovy. Pro zvolení vhodného principu zálohování bylo nejprve nutné zanalyzovat časovou náročnost a náročnost na uložení jednotlivých principů zálohování. Toho bylo dosaženo použitím vícekritériální analýzy variant. Tato analýza zobrazuje užitek jednotlivých principů zálohování na základě doby trvání, požadovaného místa na disku, obnovitelnosti a dostupnosti, díky čemuž je možné se rozhodnout, jaký princip bude použit. Následně bylo otestováno řešení strategie, které bylo zhodnoceno. Dalším krokem bylo navrženo vylepšení, které by bylo možné uplatnit při použití v praxi pro dosažení vhodnějšího řešení. V posledním kroku bylo řešení zobecněno pro použití nejen v testovacím prostředí Oracle Database, ale také v ostatních databázových systémech a případně i jiných operačních systémech.

Toto řešení je jedním z mnoha způsobů, jakými lze navrhnout strategii zálohování a obnovy a její samotnou implementaci. Jiná řešení se liší frekvencí zálohování, umístění výsledných zálohových souborů, zápisem skriptů a parametrů spouštění zálohování. Rozdílným řešením je zrcadlení celé databáze, kdy je do procesu zálohování vložen další prvek, kterým je stand-by databáze.

Jisté omezení vzniká v případě navrhovaného vylepšení na přesun katalogové databáze. Toto vylepšení není vhodné pro malé společnosti kvůli vzniku vysokých nákladů na provoz katalogového databázového serveru.

Hlavními přínosy této bakalářské práce je analýza optimálních principů zálohování, definování postupů při vytváření strategie zálohování a obnovy a definování požadavků kladených na tuto strategii. Dílčími přínosy je sjednocení požadavků na zálohování a příklad užití vybraného principu zálohování a obnovy.

Zavedením optimální strategie zálohování a pomocí její správné implementace je extrémně sníženo riziko ztráty dat. Z dlouhodobého hlediska je velmi výhodné investovat do zavedení této strategie a její údržby. Údržba zálohování je ve většině případech méně nákladná a náročná než nutnost znovu získávat data a vynaložit mnohem vyšší prostředky na získání těchto dat. V některých případech ovšem může být nemožné data znovu získat.

7 Seznam použitých zdrojů

1. CYRAN, Michele. *Oracle® Database Concepts 10g Release 2* [online]. 2005. 542 s. (PDF) [cit. 2016-09-05]. B14220-02. Dostupné z: https://docs.oracle.com/cd/B19306_01/server.102/b14220.pdf
2. CARPENTER, Tom. *Microsoft SQL server 2012 administration: real-world skills for mcsa certification and beyond*. Indianapolis, Ind.: John Wiley, 2013. 899 s. ISBN 978-1-118-48716-7.
3. FREEMAN, Robert G., PACK Charles A., STUNS Doug, BUTERBAUGH Tim. *OCF Oracle database 11g administrator certified professional study guide*. Indiana, Ind: Wiley Publishing, 2009. 838 s. ISBN 978-0-470-39513-4.
4. ASHDOWN, Lance, KYTE, Tom. *Oracle Database®: Concepts, 11g Release 2* [online]. 2015. 472 s. (PDF) [cit. 2016-09-07]. E40540-04. Dostupné z: https://docs.oracle.com/cd/E11882_01/server.112/e40540.pdf
5. MICROSOFT. *Backup and restore of SQL server databases* [online]. 2012. 288 s. (PDF) [cit. 2016-07-31]. Dostupné z: <http://download.microsoft.com/download/0/F/B/0FBFAA46-2BFD-478F-8E56-7BF3C672DF9D/Backup%20and%20Restore%20of%20SQL%20Server%20Databases.pdf>
6. BRADÁČ, Richard. *Zálohování a obnova dat – část I*. Technická příloha časopisu LOGIN. 2008, 4., č.2. str. 1-4.
7. VEČERKA, Arnošt. *Kompresce dat*. Olomouc: Univerzita Palackého, Přírodovědecká fakulta, Katedra informatiky, 2008, 65 s.
8. MATISHAK, Deirdre, FULLER, Mark. *Oracle Database 11g: Administration Workshop I*. 2nd ed. 2009. 488 s. D50102GC20.
9. DOSEDĚL, Tomáš. *Počítačová bezpečnost a ochrana dat*. Brno: Computer Press, 2004. 190 s. ISBN 80-251-0106-1.
10. FREEMAN Robert G., HART, Matthew. *Oracle RMAN 11g backup and recovery*. New York: McGraw-Hill Companies, 2010. 656 s. ISBN 978-0-07-162861-7.
11. ASHDOWN. Lance. *Database Backup and Recovery User's Guide* [online]. 2015. 604 s. (PDF). [cit. 2016-09-15]. E10642-08. Dostupné z: http://docs.oracle.com/cd/E11882_01/backup.112/e10642/toc.htm
12. VALENTA, Michal, POKORNÝ, Jaroslav. *Databázové systémy*. Praha: České vysoké učení technické v Praze, 2013. 265 s. ISBN 978-80-01-05212-9.

8 Přílohy

Příloha č. 1 - Výpis RMAN z vytvořené plné zálohy	52
Příloha č. 2 – Výpis RMAN z vytvořené přírůstkové zálohy	53
Příloha č. 3 – Výpis RMAN z vytvořené zálohy archivních záznamů	54
Příloha č. 4 – Validace plné zálohy	55
Příloha č. 5 – Validace přírůstkových záloh	56
Příloha č. 6 – Validace zálohy archivních záznamů	57
Příloha č. 7 – Výpis RMAN po obnovení databáze.....	58
Příloha č. 8 – Výpis RMAN po zotavení dat v databázi.....	59

Příloha č. 1 - Výpis RMAN z vytvořené plné zálohy

BS Key	Type	LV	Size	Device	Type	Elapsed Time	Completion Time
3564	Incr	0	5.88G	DISK		00:05:29	24-FEB-17
BP Key: 3568 Status: AVAILABLE Compressed: NO Tag:							
TAG20170224T153214							
Piece Name:							
/u04/backup/BAP/backupset/2017_02_24/HOT0_BAP_936804734_52_1_1.bkp							
List of Datafiles in backup set 3564							
File	LV	Type	Ckp	SCN	Ckp Time	Name	
1	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/system01.dbf	
2	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/sysaux01.dbf	
3	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/undotbs01.dbf	
4	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/users01.dbf	
5	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_01.dbf	
6	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_02.dbf	
7	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_03.dbf	
8	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_04.dbf	
9	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_05.dbf	
10	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_06.dbf	
11	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_07.dbf	
12	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_08.dbf	
13	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_09.dbf	
14	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_10.dbf	
15	0	Incr	1080488		24-FEB-17	/u04/oradata/BAP/BPTBS_11.dbf	

Zdroj: Autor

Příloha č. 2 – Výpis RMAN z vytvořené přírůstkové zálohy

BS Key	Type	LV	Size	Device	Type	Elapsed Time	Completion Time
3956	Incr	1	107.02M	DISK		00:01:14	24-FEB-17
BP Key: 3960 Status: AVAILABLE Compressed: NO Tag:							
TAG20170224T162312							
Piece Name:							
/u04/backup/BAP/backupset/2017_02_24/HOT1_BAP_936807792_63_1_1.bkp							
List of Datafiles in backup set 3956							
File	LV	Type	Ckp	SCN	Ckp Time	Name	
1	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/system01.dbf	
2	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/sysaux01.dbf	
3	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/undotbs01.dbf	
4	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/users01.dbf	
5	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_01.dbf	
6	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_02.dbf	
7	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_03.dbf	
8	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_04.dbf	
9	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_05.dbf	
10	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_06.dbf	
11	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_07.dbf	
12	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_08.dbf	
13	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_09.dbf	
14	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_10.dbf	
15	1	Incr	1084066		24-FEB-17	/u04/oradata/BAP/BPTBS_11.dbf	

Zdroj: Autor

Příloha č. 3 – Výpis RMAN z vytvořené zálohy archivních záznamů

BS Key	Size	Device Type	Elapsed Time	Completion Time	
-----	-----	-----	-----	-----	
3637	143.39M	DISK	00:00:02	24-FEB-17	
BP Key: 3640 Status: AVAILABLE Compressed: NO Tag:					
TAG20170224T155036					
Piece Name:					
/u04/backup/BAP/arc_back/2017_02_24/arc_936805836_54_1.bkp					
List of Archived Logs in backup set 3637					
Thrd	Seq	Low SCN	Low Time	Next SCN	Next Time
-----	-----	-----	-----	-----	-----
1	820	1069285	22-FEB-17	1081331	24-FEB-17
1	821	1081331	24-FEB-17	1081620	24-FEB-17
1	822	1081620	24-FEB-17	1081890	24-FEB-17
1	823	1081890	24-FEB-17	1082296	24-FEB-17
1	824	1082296	24-FEB-17	1082309	24-FEB-17

Zdroj: Autor

Příloha č. 4 – Validace plné zálohy

```

RMAN> validate backupset 3564;

Starting validate at 24-FEB-17
allocated channel: ORA_DISK_1
channel ORA_DISK_1: SID=40 device type=DISK
channel ORA_DISK_1: starting validation of datafile backup set
channel ORA_DISK_1: reading from backup piece
/u04/backup/BAP/backupset/2017_02_24/HOT0_BAP_936804734_52_1_1.bkp
channel ORA_DISK_1: piece
handle=/u04/backup/BAP/backupset/2017_02_24/HOT0_BAP_936804734_52_1_1
.bkp tag=TAG20170224T153214
channel ORA_DISK_1: restored backup piece 1
channel ORA_DISK_1: validation complete, elapsed time: 00:00:55
Finished validate at 24-FEB-17

```

Zdroj: Autor

Příloha č. 5 – Validace přírůstkových záloh

```

RMAN> validate backupset 3956;

Starting validate at 24-FEB-17
using channel ORA_DISK_1
channel ORA_DISK_1: starting validation of datafile backup set
channel ORA_DISK_1: reading from backup piece
/u04/backup/BAP/backupset/2017_02_24/HOT1_BAP_936807792_63_1_1.bkp
channel ORA_DISK_1: piece
handle=/u04/backup/BAP/backupset/2017_02_24/HOT1_BAP_936807792_63_1_1
.bkp tag=TAG20170224T162312
channel ORA_DISK_1: restored backup piece 1
channel ORA_DISK_1: validation complete, elapsed time: 00:00:01
Finished validate at 24-FEB-17

```

Zdroj: Autor

Příloha č. 6 – Validace zálohy archivních záznamů

```

RMAN> validate backupset 3637;

Starting validate at 24-FEB-17
using channel ORA_DISK_1
channel ORA_DISK_1: starting validation of archived log backup set
channel ORA_DISK_1: reading from backup piece
/u04/backup/BAP/arc_back/2017_02_24/arc_936805836_54_1.bkp
channel ORA_DISK_1: piece
handle=/u04/backup/BAP/arc_back/2017_02_24/arc_936805836_54_1.bkp
tag=TAG20170224T155036
channel ORA_DISK_1: restored backup piece 1
channel ORA_DISK_1: validation complete, elapsed time: 00:00:03
Finished validate at 24-FEB-17

```

Zdroj: Autor

Příloha č. 7 – Výpis RMAN po obnovení databáze

```
RMAN> restore controlfile from
'/u04/backup/BAP/backupset/2017_02_24/HOT1_BAP_936813282_71_1_1.bkp';

Starting restore at 05-MAR-17
allocated channel: ORA_DISK_1
channel ORA_DISK_1: SID=19 device type=DISK

channel ORA_DISK_1: restoring control file
channel ORA_DISK_1: restore complete, elapsed time: 00:00:03
output file name=/u04/oradata/BAP/BAP_control01.ctl
output file name=/u04/fra/BAP/BAP_control02.ctl
output file name=/u04/backup/BAP/BAP_control03.ctl
Finished restore at 05-MAR-17

RMAN> restore database from tag 'TAG20170224T153214';

Starting restore at 05-MAR-17
using channel ORA_DISK_1

channel ORA_DISK_1: starting datafile backup set restore
channel ORA_DISK_1: specifying datafile(s) to restore from backup set
channel ORA_DISK_1: restoring datafile 00001 to
/u04/oradata/BAP/system01.dbf
...
channel ORA_DISK_1: restoring datafile 00015 to
/u04/oradata/BAP/BPTBS_11.dbf
channel ORA_DISK_1: reading from backup piece
/u04/backup/BAP/backupset/2017_02_24/HOT0_BAP_936804734_52_1_1.bkp
channel ORA_DISK_1: piece
handle=/u04/backup/BAP/backupset/2017_02_24/HOT0_BAP_936804734_52_1_1
.bkp tag=TAG20170224T153214
channel ORA_DISK_1: restored backup piece 1
channel ORA_DISK_1: restore complete, elapsed time: 00:05:05
Finished restore at 05-MAR-17
```

Zdroj: Autor

Příloha č. 8 – Výpis RMAN po zotavení dat v databázi

```
RMAN> recover database until sequence 827;

Starting recover at 05-MAR-17
using channel ORA_DISK_1
channel ORA_DISK_1: starting incremental datafile backup set restore
channel ORA_DISK_1: specifying datafile(s) to restore from backup set
destination for restore of datafile 00001:
/u04/oradata/BAP/system01.dbf
...
destination for restore of datafile 00015:
/u04/oradata/BAP/BPTBS_11.dbf
channel ORA_DISK_1: reading from backup piece
/u04/backup/BAP/backupset/2017_02_24/HOT1_BAP_936811984_67_1_1.bkp
channel ORA_DISK_1: piece
handle=/u04/backup/BAP/backupset/2017_02_24/HOT1_BAP_936811984_67_1_1
.bkp tag=TAG20170224T173304
channel ORA_DISK_1: restored backup piece 1
channel ORA_DISK_1: restore complete, elapsed time: 00:00:07

starting media recovery

channel ORA_DISK_1: starting archived log restore to default
destination
channel ORA_DISK_1: restoring archived log
archived log thread=1 sequence=820
...
channel ORA_DISK_1: restoring archived log
archived log thread=1 sequence=824
channel ORA_DISK_1: reading from backup piece
/u04/backup/BAP/arc_back/2017_02_24/arc_936805836_54_1.bkp
channel ORA_DISK_1: piece
handle=/u04/backup/BAP/arc_back/2017_02_24/arc_936805836_54_1.bkp
tag=TAG20170224T155036
channel ORA_DISK_1: restored backup piece 1
channel ORA_DISK_1: restore complete, elapsed time: 00:00:15
archived log file
name=/u04/fra/BAP/archivelog/2017_03_05/o1_mf_1_820_dcrn9m7c_.arc
thread=1 sequence=820
channel default: deleting archived log(s)
...
archived log file
name=/u04/fra/BAP/archivelog/2017_03_05/o1_mf_1_824_dcrn9mb2_.arc
RECID=776 STAMP=937852739
channel ORA_DISK_1: starting archived log restore to default
destination
channel ORA_DISK_1: restoring archived log
archived log thread=1 sequence=825
```

```
channel ORA_DISK_1: restoring archived log
archived log thread=1 sequence=826
channel ORA_DISK_1: reading from backup piece
/u04/backup/BAP/arc_back/2017_02_24/arc_936812912_69_1.bkp
channel ORA_DISK_1: piece
handle=/u04/backup/BAP/arc_back/2017_02_24/arc_936812912_69_1.bkp
tag=TAG20170224T174832
channel ORA_DISK_1: restored backup piece 1
channel ORA_DISK_1: restore complete, elapsed time: 00:00:03
archived log file
name=/u04/fra/BAP/archivelog/2017_03_05/o1_mf_1_825_dcrmb2r7_.arc
thread=1 sequence=825
channel default: deleting archived log(s)
archived log file
name=/u04/fra/BAP/archivelog/2017_03_05/o1_mf_1_825_dcrmb2r7_.arc
RECID=782 STAMP=937852757
archived log file
name=/u04/fra/BAP/archivelog/2017_03_05/o1_mf_1_826_dcrmb2s7_.arc
thread=1 sequence=826
channel default: deleting archived log(s)
archived log file
name=/u04/fra/BAP/archivelog/2017_03_05/o1_mf_1_826_dcrmb2s7_.arc
RECID=781 STAMP=937852755
media recovery complete, elapsed time: 00:00:01
Finished recover at 05-MAR-17
```

Zdroj: Autor