

Ing. Magdaléna Petříková
ÚSIP Bratislava

SYSTEM RIADENIA BÁZY DÁT IDMS

1. Úvod

Pri konvenčnom spôsobe programovania jednotlivé programy musia zohľadňovať takmer všetky fyzické vlastnosti súborov, s ktorými pracujú. Toto vedie potom k značnej závislosti programov a súborov a nie je možné meniť súbor bez toho, aby to nemalo dopad aj na program a opačne. Znamená to, že je veľmi ťažké rozšíriť súbor o dáta, ktoré sú potrebné v nových aplikáciách. Vytvárajú sa nové súbory s redundantnými dátami, čo vedie k plytvaniu pamäťou a k nebezpečeniu možných rozdielov v obsahu redundantne uložených políciak. Okrem ďalších problémov, ktoré sú všeobecne známe, je tiež veľmi ťažké vytvárať sumérne informácie z dát, uložených vo viacerých súboroch. Tieto problémy odstraňuje databankový prístup, čo sa plne potvrdzuje doterajšími skúsenosťami užívateľov systémov riadenia bázy dát /SRBD/.

Nástupom jednotného systému elektronických počítačov /JSEP/ pri budovaní ASR-ov je vytvorená možnosť použiť tieto nové technológie spracovania dát i u nás. Štandardný systém banky dát pre JSEP, ktorý je vyvíjaný fy ROBOTRON v NDR pod názvom DBS/R, nie je ešte užívateľom k dispozícii. Preto užívatelia riešia tento problém aplikáciami iných SRBD. Po rozsiahlej a zodpovedajúcej prieskumnej skúsenosti rôznych databankových systémov bol navrhnutý pre úlohy nášho dátového systému riadenia bázy

dát IDMS /Integrated Data Management System/, ktorý je produkcom firmy Cullinane, USA. Systém bol v novembri 1976 skúšobne nainštalovaný na počítači EC 1030 v ústredí ústavu v Bratislave pod operačným systémom OS. Tým bola overená prenositeľnosť systému a overené na štandardnom testovacom príklade. Neskôr bol systém overovaný na vlastných testovacích príkladoch v jazyku COBOL a PL/1 na počítačoch EC 1030 a EC 1040. "Load-knižnica" systému, vytvorená na 7,25 MB diskoch zaberá 33 cylindrov. Pôvodná inštalácia systému verzie 4.0 bola v decembri 1977 aktualizovaná novšou verziou 4.5, ktorá poskytuje užívateľovi viacej možností. Táto verziu máme nainštalovanú už na počítači EC 1040 /inštalácia trvá cca 25 min/ a rozsah knižnice na 29 MB diskoch je 15 cylindrov /z toho 1 cylinder je ponechaný ako rezerva/. Vlastná inštalácia pozostáva z 2 krokov, kde v 1.kroku sa z inštaláčnej pásky prenášajú do Load-knižnice jednotlivé load-moduly a 2.krok je štandardný testovací príklad.

2. Základné charakteristiky IDMS

2.1. Štandardizácia systému

Ako ucelený programový systém riadenia bázy dát bol vypracovaný v súlade so štandardmi a špecifikáciami skupiny DBTG-COBASYL /Data Base Task Group - Conference of Data System Languages/ z roku 1971. Z toho už vyplýva, že systém IDMS poskytuje užívateľovi nasledujúce základné vlastnosti:

- ukladanie, výber a manipulácia s dátami
- oddelenie popisu dát od manipulácie s dátami
- možnosť vytvárania univerzálnej sieťovej štruktúry.

Systém ďalej poskytuje:

- Interface na ľubovoľný programovací jazyk, ktorý využíva CALL príkazy alebo jeho ekvivalenty /napr. PL/1, COBOL, FORTRAN, ASSEMBLER/.
- Možnosť použitia na JSEP a IBM 360/370 pod operačnými systémami OS, OS/VS, DOS alebo DOS/VS. /Existujú tiež verzie pre počítače SIEMENS, UNIVAC a PDP/.

- Možnosť práce buď v lokálnom režime alebo pod centrálnou verziou /viaceré aplikačné programy prístupujú k dátam do bázy/.
 - Interface na telekomunikačné monitory CICS, TASKMASTER, SHADOW II, INTERCOMM.
 - Okrem použitia týchto telekomunikačných monitorov, výrobca systému IDMS dopracoval k DB-časti /data base/ i časť telekomunikačnú - DC /data communication/ pod názvom IDMS-DC. Použitie tohto telekomunikačného monitora je viazané na verziu systému 5.0.
 - Systém od verzie 4.5 a vyššie môže pracovať v súčinnosti s vlastným generátorom výstupných zostáv IDMS/CULPRIT /verzie 4.3 CULPRIT/.
 - Súbor Data directory systému IDMS od verzie 4.5 /viď ďalej/ plní i úlohu integrovaného slovníka dát Data Dictionary. V spolupráci s kompilátorom DDDL tvorí integrovaný slovník dát /IDD - Integrated data dictionary/, účinný racionalizačný prostriedok pre správu bázy dát a ostatných užívateľov systému.
 - Umožňuje koncovému užívateľovi od verzie systému 4.5 komunikovať s bazou dát v interaktívnom režime pomocou QUERY dopytovacieho jazyka.
- Z uvedených možností, ktoré systém IDMS poskytuje, sa dá povedať, že je to ucelený integrovaný systém, ktorý užívateľovi poskytuje rad racionalizačných prostriedkov pre budovanie ASR-u.

2.2 Štruktúra bázy dát

Nezávislosť datanového obsahu od konkrétnych aplikačných programov sa realizuje formou dvoch druhov štruktúr bázy dát: - "logická štruktúra" bázy dát
- "fyzická štruktúra" bázy dát

Logická štruktúra:

- pre vytváranie vätkých logických väzieb medzi dátami sa používajú logické "sety", realizované sarnikami. To umožňuje vytvárať vätky hierarchické štruktúry dát až

po sieťovú štruktúru bez omešzovania, či už ide o počet úrovní alebo záznameov /RECORDS/. Aplikačné programy prístupujú k dátam a taktiež s nimi pracujú na úrovni vetý, alebo len jej položky v rámci setu.

Fyzická štruktúra:

- v rámci tejto štruktúry IDMS umožňuje ukladať vetý jednou z nasledujúcich metód:
 - CALCCULATED /CALC/
 - metóda ukladá vetý do BD podľa klúča, ktorý je reprezentovaný niektorou položkou vo vete. Umiestnenie vetý v rámci oblasti pre daný druh vetý sa uskutočňuje randomizačnou rutínou, ktorá pre výpočet čísla strany používa hodnotu klúča. Na vypočítanú stranu sa potom daná veta fyzicky uloží.
 - VIA
 - touto metódou sa ukladajú tie vetý, ku ktorým sa bude prístupovať ako k členovi určitého setu. Vetý sa budú ukladať fyzicky čo najbližšie k vlastníkovi SET-u.
 - DIRECT
 - táto metóda umožňuje užívateľovi kontrolu nad logickým umiestnením vetý v rámci oblasti. Užívateľ určí sám databázový klúč predtým, než sa veta do BD uloží.
 - SEQUENTIAL
 - metóda ukladá vetý vo fyzickej sekvencii podľa hodnoty určitej klúčovej položky. Táto metóda sa viaže na použitie tzv. sekvenčného módu SPF/the Sequential processing facility/, ktorý sa používa nielen pre fyzicky sekvenčné uloženie viet, ale i v prípade prístupu k vetám podľa sekundárnych alebo generických klúčov, reprezentovaných hociktorou položkou vetý.

Tieto možnosti uloženia vetý umožňujú určiť vstupné body do BD /CALC-uloženie/, ďalej prístup k vetám po logickej reťazi SET-u /VIA-uloženie/, alebo možnosť prístupu k vetám pomocou sekundárnych alebo generických klúčov /SPF/ a tým riadiť efektivnosť prístupu k dátam, uložených v BD.

2.3 Časti systému

System IDMS je ucelený programový systém riadenia bázy dát, ktorý sa skladá z niekoľkých častí. Najdôležitejšie z nich sú ďalej popísané:

2.3.1 Systémový súbor Data Directory

Všetky kompilátory IDMS pracujú v súčinnosti so súborom IDMS Data Directory /viď obr. č.1/. Tento súbor, ktorý je sám o sebe bázou dát pre systém IDMS, predstavuje definičný zdroj, do ktorého IDMS ukladá všetky informácie, ktoré sa týkajú užívateľskej bázy dát. Prvé informácie, ktoré sa do tohto súboru ukladajú z kompilácie SCHEMY, sa ďalej dopĺňajú o údaje z DMCL a SUBSCHEMA-kompilátorov, zo služobných programov /program CLUC/. Takto potom všetky časti systému komunikujú s Directory súborom a vyvolávajú ho ako zdroj charakteristík, ktoré sa dotýkajú bázy dát užívateľa. Tieto informácie možno vo forme zostáv na požiadanie vytlačiť. Veľkosť súboru Data Directory sa volí podľa požiadaviek pri inštalovaní systému. Pre zvolenú veľkosť strany 92D 8 zaberá Data Directory na 29,17 MB disku cca 7 cylindrov.

2.3.2 Jazyky a kompilátory pre popis dát

Pre popis dát slúži IDMS DDL - jazyk. Skladá sa z nasledujúcich troch častí:

- SCHEMA DDL /Data Definition Language/

- jazyk pre kompletné popísanie všetkých datových prvkov, väzieb a oblastí v rámci celkovej schémy bázy dát. Príkazy SCHEMY sa spracovávajú IDMS-kompilátorom SCHEMY, ktorý preveruje syntaxovú správnosť užívateľského popisu a ukladá informácie o takto popísanej báze dát do systémového súboru Data Directory. Vytvorí okrem výpisu vstupných príkazov /source SCHEMA listing/ i diagnostický výpis. Jazyk pre popis SCHEMY zodpovedá syntaxi Cobolu.

- DMCL /Device-Media-Control Language/

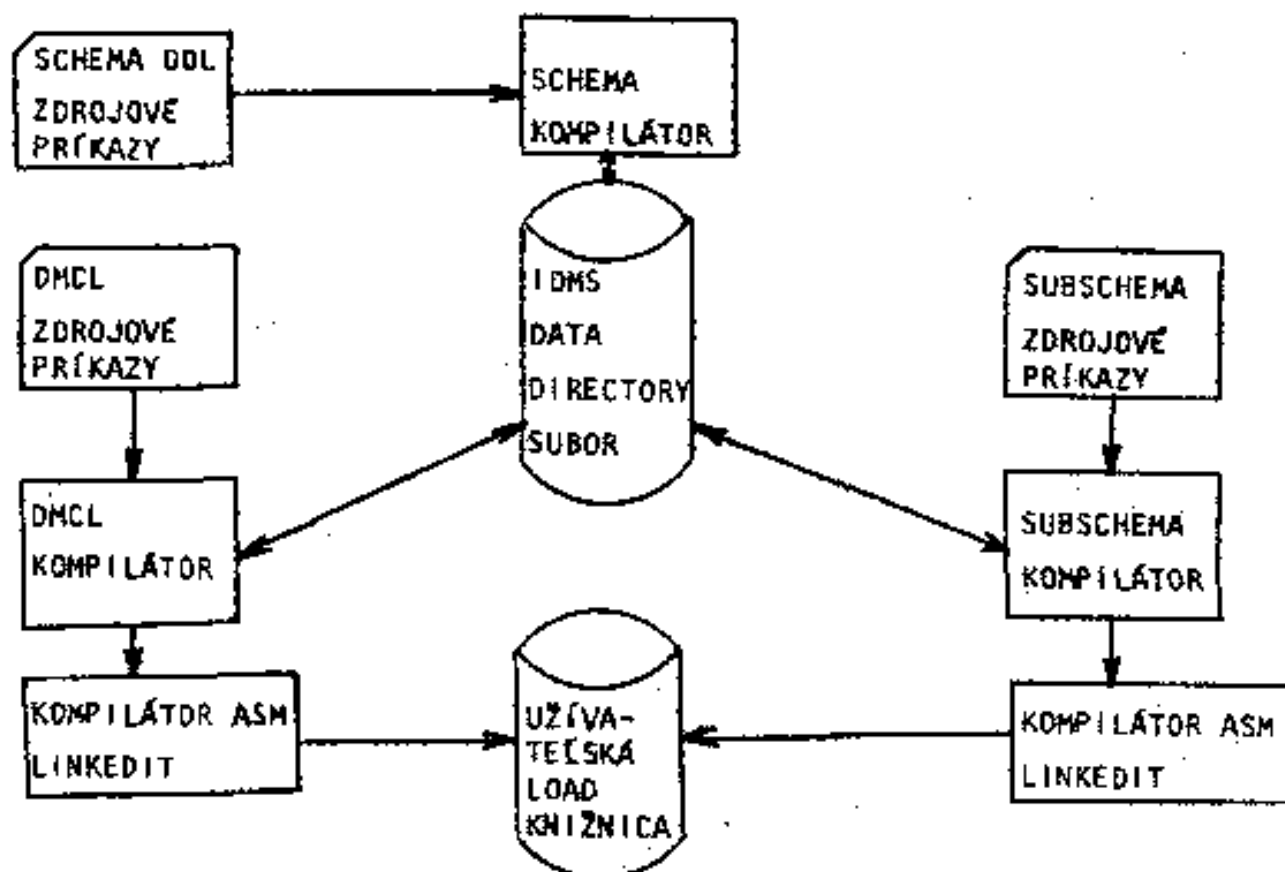
- jazyk pre riadenie rozloženia bázy dát na pamäťových médiách. Pomocou DMCL-modulov sa na základe časových,

- priestorových a operačných požiadaviek fyzicky špeci-
kujú tie databázové diskové súbory, ktoré budú v urči-
tom čase v konkrétnych podmienkach založené. Napr. tie
logické oblasti, ktoré sa budú často aktualizovať, tre-
ba priradiť do spoločného fyzického súboru. Na druhej
strane tie oblasti, ktoré sa budú menej často aktuali-
zovať, treba zahrnúť do iných fyzických súborov. Tak-
tiež je možné zahrnúť do spoločného fyzického súbo-
ru tých oblastí, ktoré sa budú často používať v urči-
tých aplikačných programoch. DML-modulov môže byť
definovaných viac.

- SUBSCHEMA - DDL

- jazyk pre definíciu subschémy. Pre jednotlivé uživa-
telské programy, alebo skupiny programov sa definujú
SUBSCHEMA - podmnožiny, ktoré predstavujú kombinácie
viac a astov, špecifické pre danú oblasť. Pre jednu
SCHEMU môže byť vytvorených viac SUBSCHEM, v závis-
losti od požadovanej počtu aplikačných "pohľadov" na
bázu dát. Všetky tieto popisy sa ako ďalšie informácie
užívateľskej bázy dát uložia pomocou SUBSCHEMA - kompilá-
tóra do Data Directory.

Vzťah medzi popísanými kompilátormi SCHEM, DML a SUBSCHEM
pre definovanie dát a systémovými súborom Data Directory je zná-
mý na obr. č.1. Data Directory používajú všetky DMS -
- kompilátory a reprezentuje zdroj, ktorý systém využíva k zis-
kaniu informácií o užívateľskej báze dát.

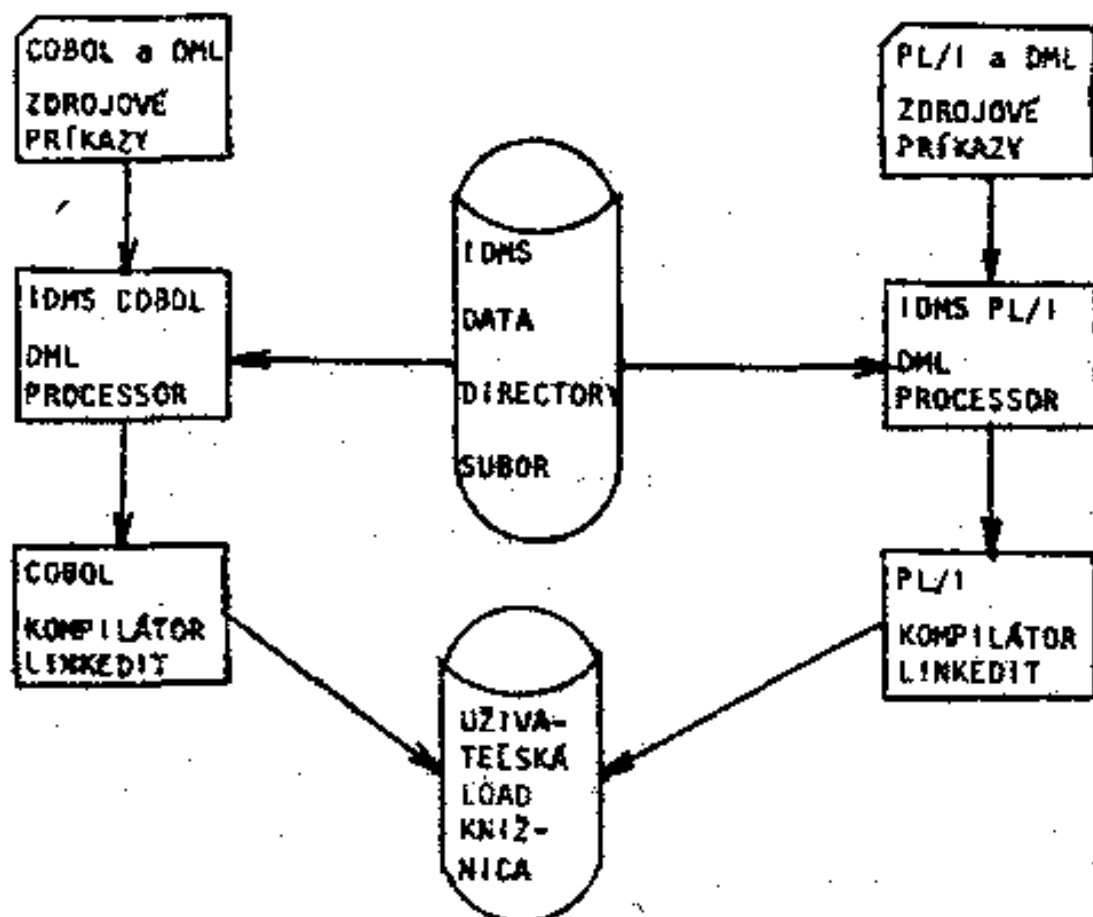


Obr. č.1 Súčinnosť systémového súboru Data Directory s kompilátormi DDL

2.3.3 Jazyk pre manipuláciu dát DML /Data Manipulation Language/

Keď sú už definované všetky dátové prvky, aplikačný programátor prístupuje k báze dát používaním DML - príkazov jazyka pre manipuláciu dát. Programátor môže písať tieto príkazy v jazyku COBOL a PL/1 priamo do zdrojového kódu tak, ako keby boli tieto rozšírenia programovacieho jazyka. Najdôležitejšie z nich sú:

- STORE:
 - pre uloženie novej vety do bázy dát /funkčne zodpovedá príkazu PL/1 WRITE/
- MODIFY:
 - pre modifikáciu vety, uloženaj v báze dát /zodpovedá funkcií REWRITE príkazu PL/1/



Obr. 6.2 IDMS-DML kompilátory. Zdrojové príkazy aplikačného programu pozostávajú z príkazov hostiteľského jazyka a príkazov IDMS-DML.

- ERASE:

- pre vymazanie /zrušenie/ vety z bázy dát

- { FIND }
- { OBTAIN }

- { vyhľadanie vety v báze dát }
- { vyhľadanie a presun do prac.oblasti }

- GET:

- presun vety do pracovnej oblasti

Prekompilátor DML tieto príkazy spracuje do CALL - príkazov, prevedie kontrolu správnosti a diagnostikou chýb a zdrojový text sa potom spracuje bežnou COBOL alebo PL/I kompiláciou. V iných programovacích jazykoch - FORTRAN, ASSEMBLER, je

možné použiť systém priamo pomocou príkazov CALL/.

Práca predkompilátorov DMLC pre COBOL a DMLP pre PL/1 v spojení s Data Directory je znázornená na obr. č.2.

2.3.4 Služobné programy /utilities/

Pre potrebnú údržbu a obnovu bázy dát poskytuje systém sadu usadených programov.

Pre vlastnú obnovu sú k dispozícii dve sekundárne kópie bázy dát, s ktorými IDMS pomocou niekoľkých služobných programov bázu dát obnovuje:

DUMP FILE

- vypísané bázu dát služobným programom IDMSDUMP na magnetickú pásku. Tým sa zachytí kompletný obraz databázy tak, ako existuje v časovom bode. Program sa riadi parametrickým štruktúrou.

JOURNAL FILE

- každej aplikáčnej programu, ktorý pracuje s bazou dát, sa priradi JOURNAL súbor na magnetickej páске, kde sa v chronologickom poradí zaznamenáva každá transakcia voči báze dát.

IDMS automaticky na tento súbor zapisuje:

a/ "BEFORE obraz"

- každé logické stránu, ktorá obsahuje požadované dáta, s ktorými bude aplikáčny program v DML - príkazoch pracovať, sa pred jej "vydania" z bázy dát do pracovnej oblasti aplikáčného programu, zapíše na journal pásku.

b/ "AFTER obraz"

- tento opis strany sa zapíše po akajkoľvek aktualizácii dát v báze, vykonanej aplikáčným programom.

c/ BEGIN checkpoint

- skončila sa aplikáčny program stáva aktívny /začiatok programu BEGIN/, je zapísaný na journal pásku kontrolný bod začiatku programu.

d/ ENDJ checkpoint

- koniec aplikáčného programu /END JOB/ sa zapíše vtedy, ak program skončí úspešne. V prípade abnormálneho konca sa nezapíše.

e/ ABRT checkpoint

- tento kontrolný bod je zapísaný pre neúspešne ukončený aplikáčny program, pracujúci pod centrálnou verziou.

S uvedenými dvomi súbormi pracujú služobné programy pri obnove bázy dát. Okrem IDMSDUMP služobného programu to sú:

- IDMSRSTR:

vykonania tohto služobného programu sa znova "reštoruje" báza dát z DUMP pásky, ktorá bola vytvorená utilitou IDMSDUMP. Program je riadený parametrickými štikami, pomocou ktorých sa systému oznamuje, aké služby, okrem znovuoživenia bázy požadujeme /možno napr. špecifikovať uloženie celej bázy dát, alebo len určitej oblasti, požadujeme výpis spod./.

- ROLLBACK:

- funkciou tejto utility je znovuoživenie "BEFORE zobrazenia" do bázy dát z journal pásky.

Program číta journal pásku smerom dozadu /back wards/ a podľa špecifikácií parametrických štikov znova uloží celú bázu dát, alebo len určité oblasti. Program takto opisuje z journal súboru "BEFORE zobrazenia" až po špecifikovaný časový okamih, ktorý sa zadá parametrickými štikami. Tým budú z bázy dát vymazané všetky úkony aplikáčnych programov, ktoré boli vykonané po tomto uvedenom časovom bode.

- IDMSRPSD:

tento prostriedok pre obnovu bázy dát je inverziou služobného programu IDMSBACK. Na rozdiel od tejto utility, ktorá znova uloží bázu dát od jej konečného stavu opísaného "BEFORE zobrazenia" až po udaný časový bod kľadu, program

IDMSRFWD znovu vybuduje bázu dát od počítačového stavu do určeného stavu s prihliadnutím na všetky zmeny, ktoré obsahuje journal pásky. Program číta journal pásku dopredu a znovu ukladá do bázy dát všetky "AFTER zobrazenia" až do špecifikovania časového bodu, ktorý sa udá parametrickým štítkom.

Okrem týchto základných služobných programov pre obnovu bázy dát je pre správu bázy dát k dispozícii ďalšia sada služobných programov.

2.4 Správa bázy dát

Pri implementácii SRSD do riešenia ASR sa stalo nevyhnutnosťou poveriť určitých pracovníkov vykonávať funkciu správy bázy dát. Pri vytvorení tejto databázovej správy si treba uvedomiť niekoľko základných aspektov:

- kapacitné nároky na plnenie tejto funkcie, ktoré sú závislé na rozsahu riešeného projektu
- vykonávanie tejto funkcie kladie značné požiadavky na kvalifikáciu pracovníkov, ktorí musia mať vedomosti o obsahovej stránke riešenia, na druhej strane tiež dostatočné vedomosti o samotnom systéme IDMS.

Správa bázy dát má zvrchovanú zodpovednosť za integritu dát v báze dát. Jej povinnosťou je nielen zodpovednosť za všetky obnovy bázy dát, journal pásky a kópie bázy dát. Je preto jedinou autoritou, ktorá môže používať spomínané služobné programy IDMS systému. Ich použitie závisí na povahe havárie, ktorá je možno špecifikovať do 3 typov:

- chyby spôsobené vykonaním užívateľského programu
- chyby spôsobené operačným systémom
- chyby na disku.

Z uvedených chýb sú najčastejšie poruchy integrity dát v báze dát z aplikačných programov. Preto je nutné, aby pred každou aktualizáciou dát v báze bol k dispozícii posledný platný DUMP súbor, alebo JOURNAL súbor pre prípad obnovy.

3. Aplikácie systému v ústave

V USIPe sa systém IDMS používa predovšetkým pre riešenie štátnej úlohy AUNUS /"Automatizované účtovnícke a hierarchicky usporiadané systémy"/, čo je jednou z dvoch nosných úloh organizácie. Cieľom úlohy je riešiť základné problémy budovania ASR-u na úrovni VHJ v celej hierarchii generálne riaditeľstvo-podnik-závod-hospodárske stredisko, a výsledky overiť na príklade konkrétneho VHJ. Pre aplikáciu bol vybraný projekt ASR VHJ Tatraske.

Systém IDMS sa tiež implementuje do riešenia ASR MP SSR, čo je druhou úlohou ústavy.

4. Záver

Vzhľadom na rozsah tohto príspevku nie je možné podrobnejšie sa zaoberať všetkými otázkami, spojenými s problematikou aplikácie SRBD - IDMS, nakoľko je to veľmi rozsiahla záležitosť. Záverom treba však spomenúť, že systémy riadenia bázy dát sú jednými z najnovších racionalizačných prostriedkov pri riešení ASR a kladú na riešiteľov značné nároky. Z našej doterajšej skúsenosti je to predovšetkým náročnosť časová, lebo zmena v technológii spracovania dát vyžaduje určitý čas na prekonanie konvenčných prístupov. Samotná práca už začína v príprave analytikov-programátorov, ktorí musia nielen navrhnúť vlastný diagram banky dát -SCHEMY /čo vyžaduje od riešiteľov mimoriadne veľkú zodpovednosť/, ale musia zvládnuť samotný systém riadenia bázy. K tomu prístupuje dobré znalosť hostujúcich jazykov spolu s jazykom riadenia práce JCL.

Vytváranie banky dát predstavuje dlhodobý interaktívny proces a v tomto zmysle je treba k tejto problematike pristupovať.

Literatúra:

- 1/ Firemná literatúra IDMS verzie 4.0 a 4.5
- 2/ Michalke P.: Koncepcia SRBD - Návrh DBTG-Codasyl informačné systémy - 1977 č.l., str. 27-38
- 3/ Miroslav Benešovec: Analýza pojmu CODASYL-SET informačné systémy - 1978 č.l., str. 25-35
- 4/ Chvalovský V.: Banky dát, ALFA 1976
- 5/ Vácha Vladislav: Funkcie a princípy databanky Socialistická akadémia ČSR, Praha 1977