

Ing. Jan Kubiček
MEZ Mohelnice, k. p.

AKTIVNÍ DOKUMENTACE DATOVÉ ZÁKLADNY

D. ANOTACE

Příspěvek popisuje jak lze jednoduchým způsobem použít výpočetní techniku pro vytvoření aktivní dokumentace o datové základně.

Na základě informací ve formě COPY-modulů /v cobolské definici/, které popisují jednotlivé datové struktury a obsahují dodatečně poznámkové příkazy je vybudován celý systém definující datovou základnu. Takový systém je nejenom schopen vydávat kompletní dokumentaci o datové základně jako je např. slovník udajů, kontrola duplicit, struktura databází, ale může být přímým zdrojem pro definici struktur v databázových systémech jako je např. IDMS, dále dovoluje provádět tisk a výpočet kapacitních nároků celé struktury.

Příspěvek uvádí popis praktické realizace takového systému na EC 1021, který byl využit jako prostředek při projektování datové základny, demonstruje výstupy z tohoto systému.

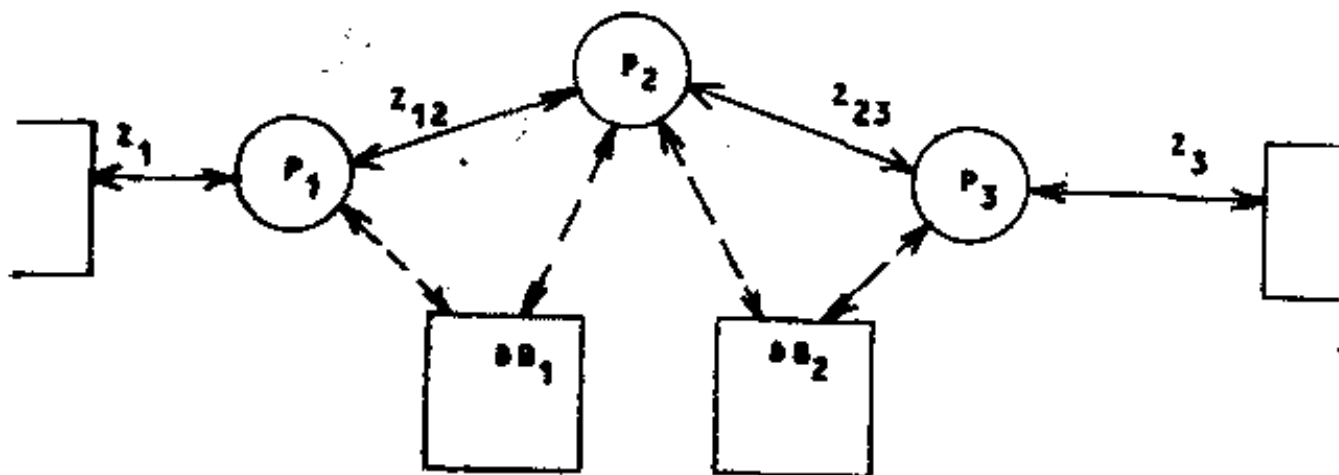
1. ŮVOD

Článek [1] odráží začátky úvah o využití výpočetní techniky při projektování ASŘP.

Při projektování rozsáhlého systému vystupují do popředí tyto otázky:

1. spolupráce různých řešitelů
2. dokumentace
3. přechod na další stupeň projekce /praktická realizace/.

Významná rovina, která je společná z různých pohledů je rovina dat. Každý subsystém popisuje svými daty část systému a nabízí tato data pro použití v jiných subsystémech. Celý systém se pak jeví jako rovina dat a rovina definovaných transformací nad těmito daty jak je znázorněno na obr. 1



Obr. 1. model systému

Zjednodušený model na obr. 1. reprezentuje syntézu mnoha modelů. Skládá se ze tří druhů elementů:

P - proces, což je asynchronní jednotka výpočtu

DB - datová báze, což je kolekce dat obsluhované systémem a updatované procesy

z - zpráva, pomocí níž procesy komunikují navzájem a s okolím systému

Popis návrhu používající tento model obsahuje:

1. Strukturální popis identifikující procesy datové báze a jejich vzájemné vztahy
2. Formální popis funkcí procesů
3. Popis obsahů datových bází

Důležitá část technických požadavků je spojena s určením struktury a formátu dat.

Data reprezentují věci a události

Struktura dat modeluje "uživatelské pochopení" věci a události

Formát dat je "okno", kterým uživatelé /a návrháři/ "vidí" věci a události.

Na návrh struktury je kladeno dále pět informačních požadavků:

1. Identifikace věci a události včetně popisu účelu v systému.
2. Vztahy mezi věcmi a událostmi.
3. Atributy věci a události.
4. Kriteria správnosti datových položek.
5. Charakteristiky použití datové struktury.

Požadavky od 1 až ad 4 zabezpečují informace pro návrh logické struktury dat. Požadavek ad 5 je nutný pro výběr vhodné fyzické struktury.

Fyzická struktura určuje umístění záznamů v paměti
Logická struktura je prostředkem určujícím nejen existenci
prvků a vztahů, ale i prostředkem pro nalezení příbuzných záznamů.

2. STRUKTURA A OBSAH DATABANKY

Technický projekt databanky, tj. struktura a předmětný obsah databanky byly vzhledem k dalšímu využívání této dokumentace zpracovány přehledně pomocí výpočetní techniky.

Na základě informací ve formě COPY-modulů, popisujících jednotlivé datové struktury, byl vybudován celý systém automatické evidence datové základny, který umožňuje zpracování kompletní dokumentace o datové základně.

Výsledky řešení jsou uvedeny v následujících kapitolách. Jedná se o:

- multiplícitní položky
- adresář položek v db
- abecední opis db vět
- struktura databází s výpočtem kapacitních nároků.

Mimo těchto uvedených funkcí je možné provést vyvoláním vhodného RUNu některé další sestavy a to adresář položek setříděný podle kódu ZSE nebo podle databází. Tyto sestavy pouze doplňují a rozšiřují možnosti systému automatické evidence datové základny.

Řešitelé jednotlivých subsystémů ASŘ produkují ve shodě s obr. 1 popis dat a popis operací nad těmito daty. Jedná-li se o systém ve kterém je 2 000 datových položek a 200 vět, pak je jistě rozumné pokusit se o počítačové zpracování tohoto "drahého" materiálu.

Důležité je najít takovou rovinu spolupráce s řešiteli jednotlivých subsystémů, aby tito byli ochotni ke spolupráci, tj. nechtít po nich víc práce. Definice výstupů od řešitelů je dána specifikací podle přílohy č. 1. Zjednodušeně řečeno: to co dřív řešitelé psali na listy popisující datové údaje, dnes zapisují do COBOLské definice. Navíc specifikují zejména údaj o počtu výskytů. Výsledkem je možnost počítačového zpracování a možnost předání definicí vět /resp. položek/ do dal-

Si úrovň projekce či realizace. Např. je možno vytvořit mag-
netickou pásku, která bude vstupem do IONS/IBD.

3. POPIS PŘÍLOH

Příloha č. 1 uvádí zkrácený metodický pokyn podle kterého
byly definovány základní stavební kameny celého systému - tj.
popis položek a vět.

Příloha č. 2 uvádí příklad skutečné definice věty.

Příloha č. 3 uvádí zkrácený výtisk struktury a výpočtu skuteč-
ného subsystému.

Literatura

[1] Veselý Karel: Automatizovaná evidence datové základny
/sborník PROGRAMOVÁNÍ 81/

1. Úvod

PŘÍLOHA č. 1

Tento metodický pokyn zavádí užití výpočetní techniky pro
dokumentaci datové základny. Na základě informací ve formě
COPY-modulů, které popisují jednotlivé datové struktury a do-
datečných poznámkových příkazů bude postaven celý systém auto-
matické evidence datové základny. Takový systém bude vydávat
kompletní dokumentaci o datové základně včetně vyšších funkcí
jako např.: slovník údajů, kontrola duplicit, struktura data-
bázi apod. Pro úspěšné splnění těchto údajů je nezbytné pečli-
vě zadávání všech informací podle následujících bodů.

2. OZNAČOVÁNÍ PRVKU DATABÁZÍ

1. Identifikátory

Databáze	DS 99
věta	RS 99
Položka	XXXXXXXX
D	označení databáze
R	označení věty

S	subsystemu
99	číslice
XXXXXX	akronym

3. DOKUMENTACE VĚT DATABÁZI

3.1 Dokumentace je provedena v rámci definice COPY-modulů.
 3.2 Používá se cobolský formulář, s definovaným polem C. Pole C začíná 21. a končí 62 kolonou včetně.
 Kolony 63 až 80 se nepoužívají.

Zápis názvů do pole C

Zápis do formulářů je zásadně od začátku příslušného pole, u dlouhých názvů je povoleno, aby začátek názvu začínal i mimo začátku pole C - pak platí pravidlo:

A	B	C	název	
11	K	111	111	
11	L	222	111222	
11	M	333	111333	
11	N	444	111333444	
11	R	555	111333555	
11	P	666	111666	Jsou povoleny maxim. 3 úrovně názvu.

3.3 Pro popis se používají:

- a/ cobolské příkazy /7. kolona je prázdná/
- b/ poznámkové příkazy /v 7. koloně je */

3.4 Cobolský příkaz - popis položky nebo skupinové položky

Pole A: úroveň položky

- 01 pouze na prvním štítku, který identifikuje zároveň i větu
- 02 se nesmí použít
- 03-49 je možno použít pro popis struktury věty, přičemž
- 03-10 jsou vyhrazeny pro vyjádření struktury
- 11-49 jsou úrovně pro popis základních položek

Pole B: identifikátor

- RS99 pro úroveň položky 01
- SXXXXXX pro položky úrovně 03-49
- identifikátor může obsahovat pouze písmena a číslice značené písmenem.

Pole C: cobolský popis položky (picture, usage, occurs, redefines na jednom řádku/.

3.5 Akronyms a názvy

se tvoří podle pravidla: nejdůležitější pojem na začátek
např. špatně - UKOLOVA MZDA , dobře - MZDA UKOLOVA

Doporučené akronymy

A	adresa	AB	absence	
C	číslo	CM	cena	CS čas
B				
D	datum	DB	doba	DE druh
E				
F				
G				
H	hodnota			
I				
J				
K				
L	lhůta			
M	množství	MZ	mzda	
N	název	NK	náklady	
O	období			
P	počet			
R	rozdíl			
S		SK	skupina	SZ sazba
T	termín			
U				
V	váha			
W				
X				
Y				
Z	způsob			

3.6 Poznámkový příkaz - dokumentační údaje

?	A	B	C
*	DB	RS99	označuje databázi, do které tato věta patří
*	PF	RS99	identifikace nadřizené věty /physical parent/ identifikace root věty
*	PC	RS99 RS99	až 10x /physical child/ nebo volnější formát:

- * PC RS99 komentář
štitků PC může být libovolný počet
a označují podřazené věty v pořadí
od leva
- * OC 999999 šti místné číslo udává odhad patřičného
počtu výskytů vzhledem k nadřazené větě.
Pro root - větu udává počet vět celé DB.
OC /OCCURENCE/ se udává s ohledem na
požadavky odhadu nároků na kapacitu disků
- tedy četnost jaká pravděpodobně bude nutná
při zpracování ON-LINE

{ * SF } [SXXXXXXXX][F]
* SFO } [L]

* SF /sequence field/ /pole seřídění/:
a/ S v této zkratce označí subsystém
b/ segment je podle tohoto pole tříděn
SF třídění vzestupné
SFO třídění sestupné
F First - první je pravidlo pro základ-
L Last - poslední ní nových výskytů vět
Toto pravidlo má význam pouze pro věty u nichž
není definováno pole seřídění není jednoznačný.
F - nový výskyt bude uložen jako první do
řady existujících /případně v rámci nejedno-
značného klíče/
L - nový výskyt bude uložen jako poslední
- standardní /nebo když štitok * SF není uveden/
d/ protože může být na štitku * SF resp SFO
nejvýše identifikátor jednoho pole, je nutné
pro některé případy použít definice skupino-
vé položky

- * S RS99 komentář /SET/
Tento štitok umožňuje definovat síťové vztahy.
RS99 identifikuje větu, která je členem SETu.

4. SCHEMA POPISU VĚTY

4.1 Popis záhlaví věty

* A B C

```

B1      RS99.
* N      NAZEY VETY
*      KOMENTAR KTERY UVADI VESKERE INFORMACE
*      SPOLECNE PRO CELOU VETU
* DB      DS99      NAZEY DATABAZE /JEM U ROOT/
{ * PP      @
  * PP      RS99 }
* OC      999999
[ * PC      RS99 RS99 RS99 RS99 RS99 RS99 RS99 RS99 RS99 RS99 ]
[ * PC      RS99      KOMENTAR ]
[ * SF      [SXXXXXXXX][F ] ]
[ * SF      [L ] ]
[ * S      RS99 ]

```

4.2 Popis_položek

4.2.1 Popis_běžné_položky /základní nebo skupinové/

```

* A      B      C
-----
11      SXXXXXXXX PIC...USAGE.....
* N      NAZEY POLOZKY
* GR      X999 - KOD ZSE
*      POPIS POLOZKY OBSAHUJE INFORMACE PRO FORMULA-
*      CI ALGORITMU. JE-LI POUZIT CISELNÍK, PAK MU-
*      SI BYT VYPSAN NEBO NA NEJ PROVEDEN ODKAZ.

```

4.2.2 Popis_položky_pro_třídění

```

* A      B      C
-----
11      S99$ORT
13      SXXXXXXXX PIC
* N      NAZEY

```

kde: { } znamená, že jedna z možností musí být uvedena
[] znamená, že údaj není povinný

Grafická úprava

Popis záhlaví věty bude od popisu položek oddělen štítkem, který má v kolonách B - 62 včetně znak - pomlčka /a v koloně 7 hvězdičku */

Zdůrazňuje se:

přesnost při popisu důležitých údajů věty, u nichž se předpokládá umístění na pevném místě ve štítku:

- a/ - definice úrovně položky musí být dvoumístná od začátku pole A
 - definice identifikátoru položky musí začínat od začátku pole B
 - cobolová definice musí začínat od začátku pole C
- b/ všechny dokumentační údaje musí mít pevný formát - zejména
 - PC
 - SF
 - OC - musí být číselné na 6 míst od začátku pole B!
 - S

5. SCHEMA POPISU DOKUMENTAČNÍCH VĚT

Fiktivní ROOT věty

5.1 K umožnění evidence celé datové základny musí každý řešitel pro všechny své subsystémy zřídít následující dvě věty.

RS1 - pokrývá ty DB, které řešitel požaduje trvale ON-LINE

RS2 - pokrývá ty věty, které se budou zpracovávat BATCH nebo věty pro archivaci.

5.2 Schema dokumentační věty

7	A ...	B	C
	01	RS1	
*	N		SUBSYSTEM ...
*	DB	ON	
*	OC	000001	
*	PC	RS99 RS99 ...RS99 RS99	10x ..
nebo			
	01	RS2	
	..		
*	DB	OFF	
*	PC	RS999 RS99 RS99...	

Štítky PC identifikují ROOT věty jednotlivých DB v příslušném subsystému.

6. NÁVRH DATABANKY

Celý návrh databanky je proveden jako hierarchický systém vět. Definice popisu vět je provedena řešiteli jednotlivých subsystémů. Řešitelé také definovali hierarchické struktury jednotlivých databází a u jednotlivých vět uvedli předpokládané četnosti výskytů.

Pomocí formálního doplnění o vrcholové věty bylo možno sdružovat všechny databáze v subsystémech a všechny subsystémy do jediného hierarchického systému. Tento systém jako jediná struktura slouží za přehlednou dokumentaci. Z popisů vět a počtu jejich výskytů je pak proveden výpočet rozsahů ve všech úrovních, takže je možné odhadnout rozsahy potřebných diskových prostorů pro implementaci jednotlivých databází, subsystémů i celého systému.

7. VÝPOČET KAPACITNÍCH NÁROKŮ

Výpočet kapacitních nároků bere v úvahu čistotu délky dat /tj. tak jak vyplývá z popisu věty/ a četnost, kterou udává řešitel v definici záhlaví věty.

Ve skutečném systému platí tato následující pravidla:

- a/ délka věty se zaokrouhluje směrem nahoru na číslo dělitelné čtyřmi,
- b/ ke každé větě budou příslušet systémové údaje /jejich délka je statisticky od 20 - 50 % celkové délky věty/.
- c/ diskovou oblast není možno naplňovat na 100 %, ale doporučuje se nepřekračovat 70 - 80 % využití diskového prostoru.

Celkově lze říci asi toto: kapacita diskového prostoru musí být 2x větší než čistý rozsah dat.

000010 01 RN01.
 000020*N KUSOVNIKOVE POLOZKY
 000030*
 000040*DB DND1 KUSOVNIKY/TP
 000050*PP 0
 000060*PC RN02 RN03 RN04 RN06 RN60 RN12 RN05
 000070*SF NCISPOL
 000080*OC 020000
 000090*S RN02 SMERNIK NA KUSOV VAZBY
 000100*-----
 000110 11 NCISPOL PIC 9(10).
 000120*N CISPOL CASTI VYSSI
 000130*GR 1015
 000140* URČUJE ČÍSLO PRÍSLUSNE ČASTI DO NÍZ
 000150* VSTUPUJE DANA ČAST (NÍZSI)
 000160*
 000170 11 NDP PIC 99.
 000180*N DRUH ČASTI
 000190*GR K018
 000200* KLASIFIKACNÍ ZNAK PRO ROZLIŠOVÁNÍ TYPU
 000210* ČASTI (FINAL, SESTAVA, PODSESTAVA, DETAIL
 000220* ATD)
 000230*
 000240 11 NNAZPOL PIC X(40).
 000250* NAZEV ČASTI ZKRACENY
 000260*GR 1082
 000270* NAZEV POLOZKY
 000280*
 000290 11 NK00P PIC 9(4) COMP-3.
 000300*N CENA KOOP POL-PLAN

01 RND SUBSYSTEM N HONA

- 175 -

01 RND SUBSYSTEM HONA ON

2	101 RND1	KUISOVNIKOVE POLOZKY	DN01 001020000	DELKA:0551	ČETNOST
1	S	SET		SET	
111	KNCISPOL	CISLO CASTI VYSEI	9(10)	N	101
111	NOP	DRUH CASTI	99	N	21
111	NNAZPOL	NAZEV CASTI ZKRACENY	X(40)	A	401
111	NKOOP	CENA KOOP POL-PLAT	9(4)	P	31

2 DELKA STRUKTURY RND1 : 071 180000 BYTES DELKA STRUKTURY VLASTNÍ DELKA

2	101 RND3	VYDEJNA	DN04 001000012	DELKA:0014	DELKA VĚTY
111	NCSTKY	CISLO STREDISKA VYDEJNY	9(5)	N	51
111	NHAB	NAKUP PRUBEZNYCH POLOZEK BEZNY NESTIC (KOB) 001000	9(6)	P	61
111	NBPPN	NAKUP PRUBEZNYCH POLOZEK OD POZATKU 001000	9(6)	P	51

3	101 RND31	STAV KN	DN04 001002500	DELKA:0048	USAGE
111	NCKN	CISLO POLOZKY KN	9(7)	N	71
111	NNHBEZ	MNOZSTVI BEZNE	9(4)	P	31
111	NKRST	KRITICKY STAV POLOZKY	9(4)	P	31
111	NSMR	SPOTREBA ZA MINULY ROK	9(4)	P	31
111	NSPBR	SPOTREBA BEZNY ROK	9(4)	P	31
112	NSPCT	SPOTREBA VE CTVRTLETI	9(4)	P	31
111	NOPPV	OBDOBÍ POSLEDNÍHO PŮHYBU-VYDEJ	9(4)	N	41
111	NOPPP	OBDOBÍ POSLEDNÍHO PŮHYBU - PRIJEM	9(4)	N	41
111	NNABR	NAKUP BEZNY ROK	04	P	31
112	NNACT	NAKUP VE CTVRTLETI	9(4)	P	31

BYTES: 000 120000

4	101 RND32	PŮHYB KN	DN06 001000001	DELKA:00191	
111	NDRP	DRUH PŮHYBU	X	A	11
111	NSTR	CISLO STREDISKA	9(5)	N	51
111	NOP	OBDOBÍ PŮHYBU	9(6)	N	61
111	NNHAP	MNOZSTVI	9(4)	P	31
111	NHHR	CARTKA PREDEPSANA K UHRADE	9(6)V99	P	41

BYTES: 000 000019

5	101 RND33	KN V TRVALEM UZIVANI PRACOVNIKA	DN04 001000004	DELKA:00241	NA JEDNU VĚTU RND31 BUDE PŮMĚRNĚ 6 VĚT RND33
111	KNCOS3	CISLO OSODNI PRACOVNIKA	9(5)	N	51
111	NNATPO	DATUM PŘEVZETI / ODEVZDANI NASTROJE	9(6)	N	61
111	NNHOP3	MNOZSTVI	99	P	21
111	NSTRED3	STREDISKO PRACOVNIKA	9(5)	N	51
111	NECH	EVIDENČNÍ ČÍSLO NASTROJE	9(7)	N	71
111	NPV	PRIJEM / VYDEJ	9	N	11

BYTES: 000 000154

3 DELKA STRUKTURY RND1 : 000 597500 BYTES = 2500(48 + 19 + 156)

2 DELKA STRUKTURY RND3 : 006 690168 BYTES

1 DELKA STRUKTURY RND1 : 000 749632 BYTES