

Věcné zadání projektu SDAT

G – Migrace dat

Obsah

1	Úvod.....	3
2	Obecné principy migrace	3
2.1	Obsah migrace MTS ==> SDAT	4
2.2	Obsah migrace MKT ==> SDAT	5
3	Struktura datových objektů aplikace MTS-ISL-SÚD-SDNS	5
4	Struktura datových objektů systému MKT	5
4.1	Struktura dat v databázi MKT	7
5	Katalog nefunkčních požadavků	9

1 Úvod

Předmětem migrace jsou data uložená v databázi stávajícího systému MTS-ISL-SÚD-SDNS (dále jen „MTS“) a vstupní zprávy (soubory) zasílané vykazujícími osobami do ČNB k následnému zpracování a data systému MKT¹ (Monitoring kapitálového trhu, dále jen „MKT“). Systémy, jejichž data podléhají migraci, jsou v dalším textu též nazývány „zdrojovými strukturami“

Z hlediska obsahu migrace lze obsah MTS rozdělit na:

- a. Data – hodnoty údajů.
- b. Metadata – metapopis, který definuje jednotlivé výkazy, registr vykazujících osob, vykazovací povinnosti.
- c. Provozní a stavové informace – zejména informace o stavu a zpracování vstupních souborů propagované až na úroveň stavů jednotlivých hodnot údajů a plnění vykazovacích povinností.

Předmětem migrace jsou rovněž data (Hodnoty údajů) za oblast MKT, která nejsou z kapacitních důvodů ukládána do databáze systému MTS. Data výkazů za tuto oblast jsou až na výjimky ukládána až do úložiště návazného systému MKT. Systém MTS tak zajišťuje pouze příjem, základní technické kontroly dat a dále zpětnou komunikaci výsledků kontrol prováděných v úložišti MKT. Po uvedení systému SDAT do provozu budou i tato data standardně ukládána do úložiště SDAT a systém MKT bude se sběrným systémem komunikovat prostřednictvím Aplikačního datového rozhraní SDAT. Předpokládaným objemem migrovaných dat z MKT jsou jednotky TB.

2 Obecné principy migrace

Pod pojmem „migrace dat“ jsou zahrnuty tyto akce a procesy, které jsou součástí dodávky SDAT:

- a) **Analýza zdrojové (MTS/MKT) a cílové struktury (SDAT)** především s důrazem na analýzu problematiky transformace dat mezi zdrojovou a cílovou strukturou.
- b) **Export objednatelům určených dat a metadat** ze zdrojových struktur a jejich import do SDAT. Dodavatel musí poskytnout SW nástroje pro export a import dat a metadat. Součástí tohoto procesu (export/import) je i kompletní zajištění nezbytně nutných transformací dat mezi zdrojovými a cílovými strukturami (tzv. „transformační mapování mezi strukturami“). Po importu musí být proveden test kompletnosti migrovaných dat a metadat, jejich správnosti a splnění integritních pravidel struktur nového SDAT.
- c) **Post-migrační** polo-automatizované nebo manuální **úpravy dat** a nastavení SDAT - lze předpokládat, že některé informace, které bude vyžadovat datový model SDAT, se ve zdrojových strukturách nevyskytují, a bude je nutné pro SDAT doplnit default hodnotami, případně realizovat určité post-migrační postupy.

¹ MKT – Monitoring kapitálového trhu, systém vyvinutý a provozovaný ČNB pro potřeby zpracování reportingu dat z kapitálových trhů.

Lze předpokládat, že procesy b) c) budou opakovány několikrát v průběhu realizace a to minimálně před každým dílčím plněním a před zahájením ověřovacího provozu. Z tohoto důvodu je nezbytné, aby byly SW nástroje pro import, export, transformační mapování a post-migrační úpravy vytvořeny tak, aby bylo možné jejich efektivní opakované spouštění.

Kromě iniciální migrace (přenesení obsahu ke zvolenému okamžiku) se tedy předpokládá i proces další inkrementální synchronizace v případě souběhu provozu systémů MTS a SDAT. V této souvislosti jsou důležité následující principy, které budou pro postupný přechod provozu z MTS do SDAT uplatněny:

1. Projektování jednoho výkazu nebude probíhat paralelně v obou systémech.
2. Sběr a zpracování výkazů bude probíhat paralelně v MTS i SDAT, přičemž ale nebude možné vykazovat současně na tutéž verzi výkazu v MTS a SDAT. Je předpoklad, že stávající výkazy budou postupně po celcích přecházet do procesu sběru v systému SDAT a to v rámci své následující verze.
3. V případě, že některý výkaz nebude přirozeně verzovat (nevznikne metodická potřeba k jeho změně), bude vytvořena „umělá“ verze, tak aby šlo aplikovat pravidlo popsané v bodě 2 výše.

Lze předpokládat, že migraci bude možno z hlediska obsahu provádět buď kompletně, nebo po definovaných věcných celcích v rámci přírůstků dodané funkcionality SW řešení. Způsob zajištění migrace dat, včetně způsobu zajištění souběhu systémů MTS a SDAT navrhne dodavatel v rámci fáze tvorby realizační studie.

Předpokládaným objemem migrovaných dat a metadat jsou v případě MTS desítky GB.

V oblasti migrace vstupních zpráv (souborů) zasílaných vykazujícími osobami je požadována plně automatizovaná migrace bez ručních nebo polo-automatizovaných zásahů.

Pokud zhotovitel bude implementovat i požadavky, které jsou v zadání označeny jako „vítané“ anebo dojde během fáze tvorby realizační studie nebo implementace systému ke změně v zadání (například bude přidána funkcionality, která dříve v zadání nebyla), musí dodavatel také splnit odpovídající požadavek na migraci údajů dané oblasti (daného požadavku).

2.1 Obsah migrace MTS ==> SDAT

1. Data (hodnoty údajů) jsou migrována v celé historii.
 - a. Není požadována migrace Hodnot údajů tzv. Odvozených údajů a jejich časových řad.
2. Metadata budou migrována v celé historii.
 - a. U kontrol platných během posledních pěti let ode dne spuštění systému SDAT je požadována transformace do jazyka SDAT, tak aby byly kontroly v systému interpretovatelné. Ostatní kontroly budou migrovány taktéž, ale není nutné zajistit jejich interpretovatelnost systémem SDAT.
 - b. Není požadována migrace tzv. Výstupních datových souborů.
 - c. Není požadována migrace tzv. Odvozených informačních prvků.
3. Provozní a stavové informace budou migrovány v celé historii v rozsahu informací pro zajištění informací o datové kvalitě a Hodnot údajů v systému SDAT.

- a. Není požadována migrace informací týkajících exportů metodik pomocí EDIFACT a systémových a uživatelských nastavení tohoto kanálu.
- b. Není požadována migrace detailů procesu zpracování jednotlivých vstupních souborů.
- c. Není požadována migrace samotných vstupních zpráv (souborů).

2.2 Obsah migrace MKT ==> SDAT

1. Data (hodnoty údajů jsou) jsou migrována z MKT v celé historii.
2. Metadata jsou standardně součástí MTS a budou migrována podle kapitoly [2.1 Obsah migrace MTS ==> SDAT](#), bod [2](#).
3. Provozní a stavové informace jsou součástí MTS a budou migrována podle kapitoly [2.1 Obsah migrace MTS ==> SDAT](#), bod [3](#).

3 Struktura datových objektů aplikace MTS-ISL-SÚD-SDNS

Podkladem pro provedení migrace ze systému MTS jsou následující technické dokumenty, které kromě definice struktur, ve kterých jsou data a metadata uložena, poskytují rovněž textový popis jednotlivých objektů a modulů systému.

Následující dokumenty jsou zpřístupňovány na základě výslovného souhlasu společnosti NESS Czech s.r.o., V Parku 2335/20, 148 00 Praha, IČO: 45786259, a slouží pouze pro účely přípravy nabídky uchazeče na veřejnou zakázku „SDAT – Sběr dat pro potřeby ČNB“ a pro účely migrace dat a jejich přípravu v rámci plnění veřejné zakázky.

1. Popis datového úložiště, viz dokument: [MtS_Migrace\MtS_SÚD v1 0.doc](#).
2. Popis objektů pro přenos a zpracování dat, viz dokument [MtS_Migrace\MtS_ISL v1 0.doc](#).
3. Popis objektů metapopisu, viz dokument [MtS_Migrace\MtS_Informační_model v1 0.doc](#).
4. Popis objektů Registru osob, viz dokument [MtS_Migrace\MtS_OM v1 0.doc](#).

Výše zmíněné dokumenty jsou uloženy v ZIP archívu „SDAT-Migrace-Přílohy.zip“, který je součástí ZIP archívu „Příloha č. 1 Smlouvy – Věcné zadání.zip/04-MigraceDatPřílohy“.

4 Struktura datových objektů systému MKT

Výkazy pro oblast MKT jsou v systému MTS organizovány ve funkční oblasti „Kapitálové trhy“. Metapopis a stavové informace související s příjmem a zpracováním těchto výkazů jsou standardně součástí systému MTS a lze je tedy migrovat stejně jako pro ostatní výkazy.

Z databáze MKT je požadována migrace hodnot údajů 14-ti výkazů s prefixem MOKAS, viz [Tabulka 1 - Výkazy MKT](#). Těchto 14 výkazů je zařazeno do 10 metodik funkční oblasti „Kapitálové trhy“, které zhruba odpovídají jednotlivým rokům (2007-2016). Hodnoty údajů budou migrovány v plném rozsahu, tj. za všechna období.

ID výkazu	Název výkazu	Platnost do	Periodicita
MOKAS40	Hlášení obchodů s investičními nástroji - OCP	1.1.4000	denní
MOKAS41	Informace o pokynech - DEOCP	1.1.4000	měsíční
MOKAS42	Informace o obchodech - DEOCP	1.1.4000	měsíční
MOKAS43	Informace o nástrojích - DEOCP	1.1.4000	měsíční
MOKAS44	Informace o osobách - DEOCP	1.1.4000	měsíční
MOKAS50	Informace o obchodech - ORT/MOS	1.1.4000	denní
MOKAS60	Informace o nástrojích přijatých k obchodování	1.1.4000	denní
MOKAS61	Informace o objednávkách v obchodním systému	1.1.4000	denní
MOKAS62	Informace o kotacích v obchodním systému	1.1.4000	denní
MOKAS63	Informace o cenách a objemech obchodů	1.1.4000	denní
MOKAS65	Informace o účastnících RT/VS	1.1.4000	denní
MOKAS70	Informace o indexech	1.1.4000	denní
MOKAS80	Informace o vypořádání transakcí	1.1.4000	denní
MOKAS82	Informace o nástrojích přijatých k vypořádání	31.12.2012	denní

Tabulka 1 - Výkazy MKT

4.1 Struktura dat v databázi MKT

Data jsou v databázi MKT ukládána do tabulek odpovídajícím jednotlivým datovým oblastem výkazů, kde sloupce představují parametry a informační prvky (ukazatele). Celkem se jedná o 17 tabulek pro výkazy s prefixem MOKAS. Struktura těchto tabulek se liší mezi jednotlivými metodikami (s prefixem „MKT“), kterých bylo pro funkční oblast Kapitálové trhy publikováno v letech 2007 – 2016 celkem 10. Popis struktury těchto tabulek je uveden v příloze [SDAT-ZD.G-MigraceDat_DeskripceMKT.xlsx](#) uložené v ZIP archívu „SDAT-Migrace-Přílohy.zip“, který je součástí ZIP archívu „Příloha č. 1 Smlouvy – Věcné zadání.zip/04-MigraceDatPřílohy“). V souboru jsou jednotlivé tabulky uvedeny na samostatných listech Excelu a metodiky v řádcích. Pro každý sloupec každé tabulky a metodiku, ve které byl tento sloupec zařazen, je uveden datový typ.

Kromě těchto sloupců obsahuje v MKT každá datová tabulka sloupec, který uchovává identifikátor Vydání výskytu výkazu, kterým byla data do ČNB doručena a další sloupec s časovým údajem předání dat ze systému MTS do MKT. V níže uvedené statistice počtu sloupců nejsou tyto dva sloupce zahrnuty. Velikost jednotlivých tabulek v MKT k 5.6.2016 v bytech a počet záznamů v nich udává [Tabulka 2 - Velikosti datových oblastí v MKT](#).

Datová oblast	Platnost do	Počet sloupců	Velikost tabulky (bytes)	Počet záznamů
MOKA40_11	1.1.4000	25	8 355 568 289	28 912 001
MOKA40_21	1.1.4000	6	11 782 870	138 622
MOKA41_11	1.1.4000	40	74 045 089 731	368 383 531
MOKA42_11	1.1.4000	30	67 100 070 816	316 509 768

MOKA43_11	1.1.4000	17	871 614 486	6 316 047
MOKA44_11	1.1.4000	13	7 183 091 632	45 752 176
MOKA50_11	1.1.4000	27	6 784 341 181	23 475 229
MOKA50_21	1.1.4000	5	638 479 795	7 511 527
MOKA60_11	1.1.4000	38	117 175 301	414 047
MOKA61_11	1.1.4000	31	29 352 636 694	151 302 251
MOKA62_11	1.1.4000	13	49 833 551 040	479 168 760
MOKA63_11	1.1.4000	28	142 855 288	830 554
MOKA65_11	1.1.4000	16	28 440 360	160 680
MOKA70_11	1.1.4000	10	4 107 950	33 950
MOKA70_21	1.1.4000	6	355 497 639	5 007 009
MOKA80_11	1.1.4000	34	13 272 351 635	43 515 907
MOKA82_11	31.12.2012	38	418 328 053	1 478 191

Tabulka 2 - Velikosti datových oblastí v MKT

Pro odhad nárůstu jednotlivých tabulek v budoucnosti je možné použít roční přírůstky v posledních letech, které od roku 2013 zachycuje [Tabulka 3 - Roční nárůsty datových oblastí v MKT](#).

Datová oblast	2013 (bytes)	2014 (bytes)	2015 (bytes)	2016 (bytes k 5.6.2016)
MOKA40_11	946 384 832	1 021 314 729	1 967 331 664	954 385 219
MOKA40_21	575 110	262 990	300 815	110 075
MOKA41_11	10 745 876 472	10 315 937 070	15 356 916 369	9 989 760 300
MOKA42_11	10 980 938 984	9 775 703 084	13 593 871 208	6 970 548 552
MOKA43_11	199 594 782	117 300 966	139 566 024	57 180 162
MOKA44_11	1 108 810 773	1 068 134 742	1 864 455 070	889 088 174
MOKA50_11	829 162 964	809 872 503	960 548 722	437 351 792
MOKA50_21	106 104 310	103 214 820	126 215 990	58 346 040
MOKA60_11	165 755 930	238 340 053	274 308 504	117 175 301
MOKA61_11	2 160 491 864	4 385 551 390	12 598 704 544	5 052 773 262
MOKA62_11	7 151 676 688	13 146 978 832	17 826 186 872	8 293 399 296
MOKA63_11	19 791 008	22 291 888	23 629 876	10 538 440
MOKA65_11	4 180 209	4 143 216	4 001 616	1 709 112
MOKA70_11	806 586	823 284	789 646	339 768
MOKA70_21	38 253 664	60 744 050	78 845 287	34 089 159
MOKA80_11	2 326 376 215	1 922 109 695	1 523 812 330	679 169 730

Tabulka 3 - Roční nárůsty datových oblastí v MKT

5 Katalog nefunkčních požadavků

ID požadavku	Název požadavku	Popis Požadavku	Důležitost	Kategorie
MIG_1.0	Migrace – opakovatelnost	<p>Je zajištěna migrace metadat i dat ze systémů MTS a MKT a tuto migraci lze provádět opakovatelně. Systém tak disponuje modulem, který umožňuje opakovaný export dat a metadat a jejich import do systému SDAT.</p> <p>V případě, že nelze data migrovat automatizovaně (v důsledku změny filozofie nového systému), pak systém umožňuje metadata/data před provedením doplnit/modifikovat (viz MIG_3.0)</p> <p>Importní modul provádí v rámci procesu importu dat předepsané byznys kontroly.</p>	Závazný	1
MIG_2.0	Migrace – Výkazy - per partes	Je umožněno, aby bylo možno migrovat metadata i data (včetně kompletní historie) za vybraný Výkaz.	Závazný	1
MIG_3.0	Migrace – doplnění/úprava hodnot	Proces migrace umožňuje doplňovat chybějící data či modifikaci existujících dat, tak aby mohla být migrace provedena.	Závazný	1