

Systémová integrace (SI) [prezentace](#)

V užším pojetí se jedná o integraci systémů (aplikací) do jednoho celku a vytvoření takové architektury informačního systému, která bude efektivně podporovat business procesy v organizaci. V širším pojetí jde o pojem pro celý proces, který je nezbytný pro efektivní fungování především rozsáhlejších informačních systémů.

Systémový integrátor

Jedná se o poskytovatele služeb systémové integrace. Sahrává významnou roli při systematizaci požadavků zadavatele na vývoj a rozvoj IS. Řídí, koordinuje a přebírá garanci za efektivní realizaci business cílů z hlediska jejich termínů, rozpočtu a kvality.

- Je garantem řádného návrhu řešení a jasné specifikace zadání na jeho realizaci.
- Je koordinátorem požadavků na vstupy.
- Je kontrolorem výstupů.
- Je zodpovědný za dodání funkčního řešení zadavateli.

Správně navržená řešení jsou obvykle představována:

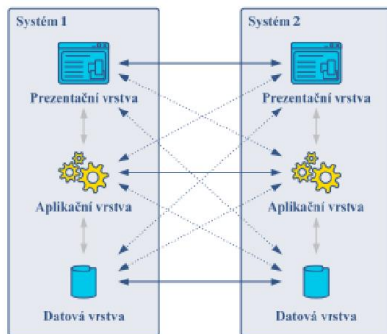
- zjednodušením stávající architektury,
- nižšími náklady na modifikaci stávajících systémů,
- nižšími náklady na implementaci a integraci nových systémů,
- automatizací business procesů (snížení nákladů, zrychlení procesu),
- snadnější integrací se systémy externích subjektů.

4 přístupy k systémové integraci

1. Replikace dat,
2. ETL (Extract Transform Load),
3. EAI (Enterprise Application Integration),
4. EII (Enterprise Information Integration).

Úrovně integrace

1. Integrace na úrovni datové (perzistentní) vrstvy,
2. Integrace na úrovni aplikační vrstvy,
3. Integrace na úrovni uživatelského rozhraní,
4. Integrace mezi rozdílnými vrstvami architektury IS.



Integrační styly

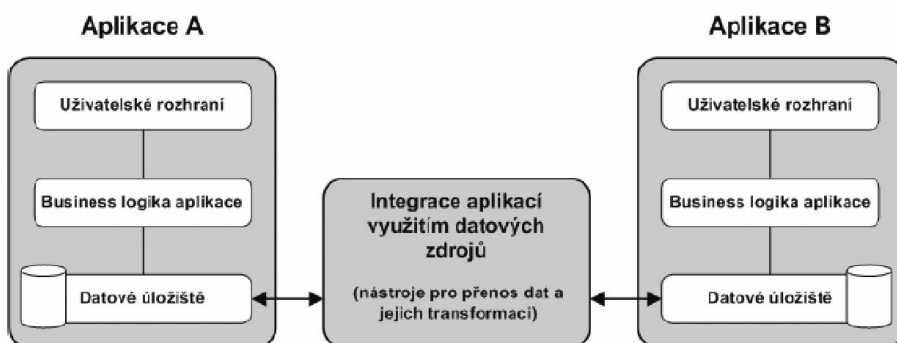
Jedná se o obecné přístupy k integraci systémů či aplikací. Představují možné způsoby komunikace mezi jednotlivými systémy.

Architektury integrace

Zabývají se otázkou, jak konkrétně budou uspořádány vazby mezi jednotlivými systémy či aplikacemi. Mezi nejpoužívanější patří:

- Point to Point
- Message Broker
- Publish/Subscribe
- Message Bus

Integrace na úrovni datové vrstvy



Většina informačních systémů a aplikací využívá pro trvalé uložení provozních dat svou vlastní databázi. V prostředí organizace, kde je využíváno více systémů s oddělenými datovými úložišti, pak často dochází ke vzniku redundancí a nekonzistence dat. Cílem integrace na této vrstvě je především eliminovat tyto nedostatky a poskytnout informačním systémům a

aplikacím přístup k datům ostatních systémů. Rovněž se může jednat o zajištění pohledu na celopodniková data pro analytické účely.

Integrace na datové vrstvě se dá realizovat několika způsoby:

- Přenos souborů
- Sdílená databáze
- Sdílené soubory
- Replikace dat
- ETL procesy

Přenos souborů

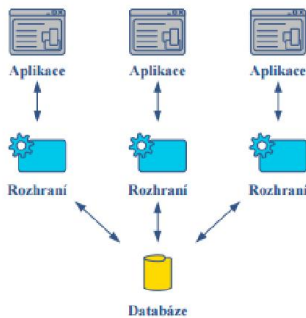
Jeden z nejjednodušších způsobů integrace na datové vrstvě je export, přenos a import souborů. Rozhraní mezi systémy je v tomto případě realizováno samotným souborem, který má určitý definovaný formát, používaný jak zdrojovou, tak i cílovou aplikací. Výhodou je zejména skutečnost, že ukládání dat do souborů podporují všechny operační systémy a pro realizaci není nutné využívat další technologie. Princip ukazuje obrázek:



U tohoto stylu integrace je problémem především zajištění unikátnosti názvů souborů a jejich odstranění v případě, kdy již nejsou potřebné. Již samotné určení, že soubor není dále potřebný, představuje problém, který je nutné v průběhu integrace vhodně vyřešit, včetně stanovení kompetencí k odstraňování souborů. Dalším problémem je pak souběžný přístup k exportovanému souboru. Rovněž je nutné vyřešit otázku, jak často bude export či import souborů probíhat a zda jednotlivé kroky budou prováděny automatizovaně nebo budou iniciovány ručně uživatelem. Nedostatkem tohoto stylu integrace pak může být neaktuálnost dat, která jsou přenášena a následně importována s určitým časovým zpožděním.

Sdílená databáze

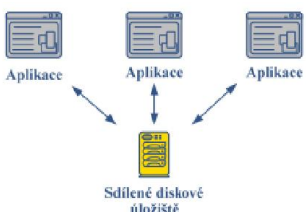
U tohoto stylu integrace používá více informačních systémů či aplikací jednu databázi pro společné ukládání provozních dat. Jednotlivé systémy tedy využívají společnou databázi a společný datový model. Výhodou oproti předchozímu stylu integrace pomocí přenosu souborů je aktuálnost dat, kdy nedochází k časovému zpoždění a veškerá data jsou vždy aktuální. Způsob integrace prostřednictvím sdílené databáze znázorňuje obrázek:



Obecně je použití sdílené databáze slabým místem celého řešení, kdy může dojít ke zpomalení běhu jednotlivých aplikací způsobenému souběžným přístupem k této společné databázi. Klíčovým problémem tohoto integračního stylu je zejména vhodně navržený datový model a následně jeho údržba. Správně navrhnout datový model tak, aby současně vyhovoval potřebám všech integrovaných aplikací, bývá obvykle podstatně obtížnější a časově náročnější úkol než v případě návrhu pro jeden samostatný systém. Datový model je pak obvykle poměrně rozsáhlý, což může přinášet problémy z hlediska výkonu. Problémy mohou nastat také u některých starších databázových technologií, které mnohdy nedostatečně podporují souběžné zpracování transakcí či případně vůbec nepodporují transakční zpracování dat.

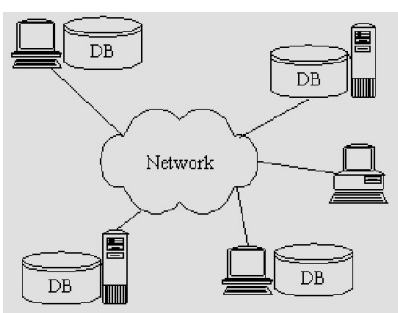
Sdílené soubory

V případě integrace pomocí sdílených souborů je architektura a princip podobný předchozímu stylu, tj. integraci pomocí sdílené databáze. Zde se však jedná o využití společného diskového úložiště a společných souborů obsahujících provozní data. Jedná se o jednoduchý způsob integrace na datové vrstvě, jehož problémem je zejména zajištění a ošetření souběžného přístupu k souborům. Princip je znázorněn na obrázku:



Replikace dat

Replikace je proces kopírování a údržby databázových objektů ve více databázích, které tvoří distribuovaný databázový systém. Změny aplikované na jednu část jsou zachyceny a uloženy lokálně před posláním a aplikováním na každou vzdálenou lokaci. Replikace poskytuje uživateli rychlý lokální přístup ke sdíleným datům a chrání dostupnost dat z aplikací, protože existují alternativní možnosti přístupu k datům. Dokonce, pokud se jedno místo stane nedostupným, uživatel může pokračovat s dotazem nebo změnou dat ve zbývajících místech.



Electronic Data Interchange (EDI)

Definice EDI

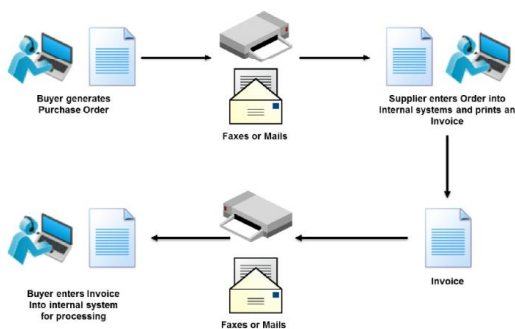
způsob komunikace mezi dvěma nezávislými subjekty při které dochází k výměně standardních strukturovaných obchodních a jiných dokumentů elektronickou formou

usnadnění pravidelné výměny strukturovaných dat mezi IS stanovením pravidel této výměny

Příklad pro rychlé vytvoření představy o EDI

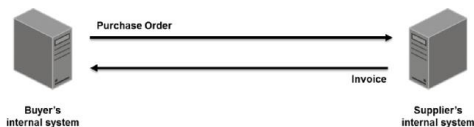
Objednání zboží bez využití EDI

1. Odběratel vytvoří ve svém IS objednávku.
2. Odběratel objednávku vytiskne.
 - o na papír či do PDF
3. Odběratel objednávku odešle.
 - o faxem či emailem
4. Dodavatel objednávku přepíše z papírové podoby do svého IS.
5. Dodavatel objednávku zpracuje, dodá a vyfakturuje.



Objednání zboží s využitím EDI

1. Odběratel vytvoří ve svém IS objednávku.
2. Odběratel objednávku vytiskne (na papír či do PDF).
3. Odběratel objednávku odešle (faxem či emailem).
4. Dodavatel objednávku přepíše z papírové podoby do svého IS.
5. IS odběratele odešle objednávku do IS dodavatele.
 - o v elektronické podobě prostřednictvím EDI
6. Dodavatel objednávku zpracuje, dodá a vyfakturuje.



Kompatibilita

- EDI bylo poprvé použito v 60. letech v automobilovém průmyslu.
- Komunikace prostřednictvím EDI je standardizována mezinárodním standardem UN/EDIFACT.
- ⇒ Nejedná se o novinku s krátkou životností, nýbrž o léty ověřenou technologii.

Standard UN/EDIFACT

- United Nations / EDI for Administration, Commerce and Transport
- spravovaný OSN
- mezinárodně uznávaný
- obecný mezioborový soubor standardů, v rámci kterého vznikají standardy pro jednotlivá odvětví:
 - o obchod (EANCOM)
 - o bankovníctví
 - o státní správa
 - o ..
- standardizuje předávání dokumentů mezi nezávislými IS

Standard EANCOM

- podmnožina UN/EDIFACT
- standardizuje elektronickou komunikaci prostřednictvím EDI v dodavatelско-odběratelských vztazích
- definuje podobu více než 50 různých typů zpráv (=dokumentů) obvykle používaných v obchodní praxi, například:

ORDERS	Objednávka
INVOIC	Faktura
INVRPT	Přehled zásob
DESADV	Avízo o odeslání zboží
RECADV	Potvrzení příjmu zboží
PRICAT	Katalog zboží a cen
CONTRL	Kontrolní zpráva
APERAK	Potvrzení o převzetí zprávy aplikací

Příklad podoby EDI dokumentu

- Dokumenty EDI obsahují stejná data, jaká byste mohli běžně najít v papírové formě dokumentu používaného pro stejný účel.

Zpráva typu INVOIC, čili *faktura*:

```
UNB+UNOB:1+PARTNER ID:ZZZ+0038977332:1:MFGB+001230:0000+000000000001++INVOIC++++1'UNH+0001+INVOIC:S:93A:UN'BGM+380+INVOICE-NBR+9'DTM+137:20000101:10
```

Zpráva typu CONTRL, respektive *potvrzení o přijetí (výše uvedené) faktury*:

```
UNB+UNOB:1+0038977332:1:MFGB+PARTNER ID:ZZZ+150426:1534+100+++++1'UNH+1+CONTRL: :ONT:UN'UCI+0000000000001+PARTNER ID:ZZZ+0038977332:1:MFGB+7'UCM+000
```

Příklady EDI dokumentů a různé jejich online generátory naleznete na <http://www.edidev.com/menuExamples.html>.

Komunikační kanály využitelné pro EDI zprávy

- Standardně jsou EDI zprávy přenášeny po sítích s přidanou hodnotou tzv. VAN (Value-Added Networks), například síť x.400.
- V roce 2002 publikovala IETF dokument RFC 3335, který nabízí standardizovaný a bezpečný způsob, jak přenášet EDI zprávy pomocí emailu.
- Od roku 2005 připravuje EDIINT (pracovní skupina IETF) podobný dokument pro FTP a HTTP přenosy.
- VAN síť a poskytovatelé EDI služeb však zůstávají.

Přínosy EDI komunikace

- zrychlení výměny/toku dokumentů
 - snižuje prodlevy v předání dokumentů
 - odpadá přepisování zprávy do systému příjemce
- standardizace podoby komunikace (vyměňovaných dokumentů) mezi rozdílnými IS
 - předávané dokumenty mají jasně definovanou strukturu
- snížení finančních nákladů s výměnou dokumentů spojených (tisk, poštovné, administrativní (personální náklady))
- nárůst kvality informací v dokumentech obsažených
- zvýšení bezpečnosti předávaných dokumentů
- zvýšení spolehlivosti
 - eliminace chyb vznikajících při ručním zadávání dat
 - eliminace problémů s (ne)čitelností faxu
- možnost větší automatizace procesů (předávané dokumenty jsou lépe automaticky zpracovatelné)
- zjednodušení archivace dokumentů
- z předchozích bodů plynoucí *zvýšení efektivity prováděných procesů*

EDI konvertor

EDI konvertor je aplikace, která mění (konvertuje) data mezi dvěma jejich podobami:

- formát poskytovatelský/srozumitelný IS (zvaný někdy jako "inhouse" formát)
 - například XML, PDF, TXT
- EDI formát (používaný pro přenos zpráv)
 - například UN/EDIFACT, XML

Konvertory jsou vysoce specializovaným druhem softwaru, díky čemuž mívají také vysoce speciální ceny. Kromě vysokých pořizovacích nákladů nutné počítat také s náklady na jejich správu, údržbu, provoz a také s náklady na jejich aktualizace, neboť se jedná o velmi dynamické odvětví.

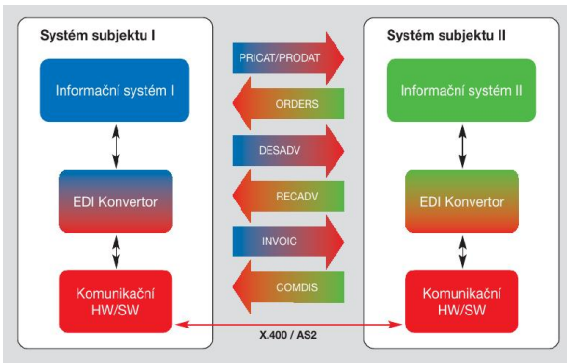
Možné podoby EDI

Podle způsobu distribuce zpráv může EDI nabývat následujících podob:

- výměna zpráv přímo mezi koncovými subjekty (klasické původní EDI řešení)
- výměna zpráv prostřednictvím VAN operátora
- výměna zpráv prostřednictvím poskytovatele EDI služeb (v současnosti nejvíce využívaná varianta)

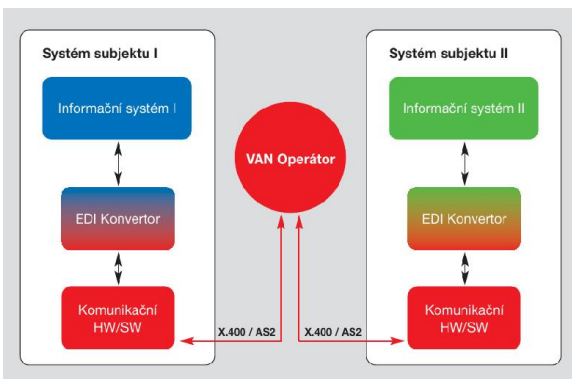
Výměna zpráv přímo mezi koncovými subjekty

- EDI komunikace je celá v režii komunikujících stran bez využití třetího subjektu.
- Nejdražší z variant.
- Každá účastník komunikace
 - vlastní EDI konvertor,
 - vlastní komunikační hardware a software pro připojení do datové sítě x.400,
 - vlastní (samozřejmě placenou) přípojku do datové sítě x.400 (standardní síť pro výměnu EDI zpráv),
 - a celé EDI řešení udržuje a aktualizuje na své vlastní náklady.



Výměna zpráv prostřednictvím VAN operátora

- VAN = Value Added Network (= síť s přidanou hodnotou)
- VAN operátor provozuje VAN a komunikaci přes ni proudící obohacuje o další služby:
 - přebírá zodpovědnost za zajištění distribuce zpráv
 - přebírá část starostí s provozem EDI řešení svého zákazníka
 - poskytuje záruku za distribuci nezkreslené informace k jejímu adresátovi
 - obvykle figuruje také jako dodavatel software (konvertoru a komunikačního software pro napojení na VAN síť)



Výměna zpráv prostřednictvím poskytovatele EDI služeb

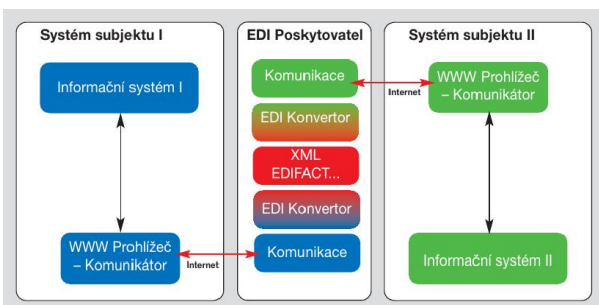
Poskytovatel EDI služeb

- je VAN operátorem (provozuje VAN), který
- je napojen na VAN síť ostatních EDI poskytovatelů,
- a provádí konverzi zpráv (provozuje EDI konvertor).

Přínosy pro uživatele:

- Uživatel nepotřebuje žádný EDI konvertor.
- ⇒ Uživatel nepotřebuje žádný specializovaný software/hardware pro komunikaci po x.400.
- ⇒ Uživatel nepotřebuje přípojku do standardní sítě pro přenos EDI zpráv (x.400).
- ⇒ Uživateli stačí běžná internetová přípojka (o propustnosti vytáčeného připojení ve většině případů).
- ⇒ Minimální pořizovací náklady, neboť celé EDI je u poskytovatele EDI.

Informační systém jednoho subjektu odešle data přes síť internet k EDI poskytovateli, který pomocí konvertorů změní data na svůj vnitřní standardizovaný formát (XML, UN/EDIFACT). Před odesláním zpráv adresátovi pak EDI konvertor změní data na formát příjemce a odešlou se pomocí internetu do informačního systému druhého subjektu.



Bezpečnost EDI komunikace

Bezpečnost dat přenášených prostřednictvím EDI je zajištěna

- šifrováním komunikačních kanálů mezi účastníky EDI komunikace,
- a implementací elektronického podpisu a šifrování přímo do zpráv EDIFACT.

Přínosy elektronického podpisu v EDI komunikaci

- identifikuje původce podpisu
 - příjemce bezpečně ví, kdo je autorem či odesílatelem zprávy
- zaručuje integritu zprávy
 - příjemce má jistotu, že zpráva nebyla změněna v průběhu transportu, což ruční podpis může zajistit jen stěží
- zaručuje nepopíratelnost
 - odesílatel nemůže popřít, že danou zprávu s daným obsahem opravdu odeslal
- nelze napodobit
 - prostředky k podepisování může mít daná osoba pod svou výhradní kontrolou

Vybrané typy EDI zpráv

zkratka	název	popis
APERAK	Application error and acknowledgement message	Zprávu zasilá příjemce EDI zpráv jejich odesílateli. Zpráva informuje o úspěšném zpracování nebo chybách při zpracování zpráv aplikací příjemce.
	Potvrzení o převzetí zprávy aplikací	
COMDIS	Commercial Dispute	Zprávu zasilá odběratel prodávajícímu pokud se vyskytne v přijaté faktuře nějaká nesrovnalost (nesprávná cena, špatná identifikace zboží, nedodané zboží apod.). Odběratel touto zprávou oznamuje odmítnutí faktury s chybami nebo její případné přijetí s výhradami a požadavek na opravu chyb.
	Obchodní námítka	
CONTRL	Syntax and Service Report	Zprávu zasilá příjemce EDI zpráv jejich odesílateli. Kontrolní zpráva syntakticky potvrzuje nebo zamítá přijatou zprávu nebo funkční skupiny zpráv.
	Kontrolní zpráva	
DESADV	Despatch Advice	Zpráva specifikuje podrobnosti o dodávaném zboží podle podmínek dohodnutých mezi kupujícím a dodavatelem. Zpráva by měla být vždy zasílána před fyzickou dodávkou zboží, aby umožnila příjemci získat přesné informace o dodávaném zboží (jeho množství, způsobu balení apod.) a tím mu umožnila včasnou přípravu na fyzickou převjemku zboží. Obdoba dodacího listu.
	Avízo o odeslání zboží	
IFTMAN	Arrival Notice	Zprávu zasilá přepravní firma příjemci zásilky s detaily o jejím dodání. Zprávu je také možno využít jako důkaz o dodání. Jedna zpráva vždy odpovídá jedné zásilce.
	Avízo příchodu zásilky	
INVOIC	Invoice	Zprávu zasilá dodavatel odběrateli jako výzvu k zaplacení za zboží či služby. Stejná zpráva může plnit i funkci proforma-faktury, zálohové faktury, dluhopisu nebo dobropisu. Prodávající může fakturovat jednu nebo více transakcí (podle jedné nebo více objednávek). Faktura může obsahovat údaje o platebních podmínkách, podrobnosti o dopravě a další doplňující informace pro celní nebo statistické účely u zahraničních zásilek.
	Faktura	
INVRPT	Inventory Report	Zpráva umožňuje odběrateli a dodavateli vyměňovat si informace o současných a plánovaných nebo cílových skladových zásobách. Obvykle obsahuje informaci o počátečním stavu, současném stavu, pohybu zásob apod.
	Přehled zásob	
ORDERS	Purchase Order	Zprávu zasilá zákazník dodavateli, aby objednal zboží nebo služby v požadovaném množství spolu s případnou specifikací místa a termínu dodání.
	Objednávka	
PARTIN	Party Information	Tato zpráva by měla být první informací, kterou si obchodní partneři vymění při zahájení EDI komunikace. Poskytuje komunikujícím organizacím základní matriční data o partnerské organizaci, jako je GLN, adresa organizace, kontaktní osoby a další data administrativního, obchodního či finančního charakteru. Zpráva je také zasílána při změně výše uvedených údajů.
	Informace o organizaci	
PRICAT	Price/Sales Catalogue	Zprávu zasilá dodavatel svým zákazníkům. Zpráva je využívána buď jako aktuální katalog či seznam všeho zboží nabízeného dodavatelem nebo jako oznámení o změnách v nabízeném sortimentu zboží. Katalog by měl obsahovat také další logistické, obchodní a cenové informace k jednotlivým uvedeným položkám.
	Katalog zboží a cen	
RECADV	Receiving Advice	Zprávou potvrzuje odběratel dodavateli příjem zboží. Zpráva je zpravidla odpovědí na zprávu DESADV a vztahuje se k jednomu expedičnímu místu dodavatele a jednomu místu dodání. Ve zprávě také mohou být řešeny případné nesrovnalosti v dodávce zboží.
	Potvrzení příjmu zboží	
RETANN	Announcement For Returns	Zpráva je využívána, pokud chce organizace vrátit jiné organizaci dodané zboží z určitého specifikovaného důvodu (k opravě, z důvodu poškození apod.) nebo pokud byly zjištěny nesrovnalosti (omylem dodané zboží, zboží s prošlou záruční lhůtou) až po potvrzení příjmu zboží zprávou RECADV a odběratel požaduje výměnu zboží apod.
	Oznámení o vrácení zboží	

Master Data Management (MDM)

MDM představuje souhrnnou sadu procesů, přístupů, metodologií a nástrojů, která centrálně definuje a spravuje kmenová data organizace (master data). Úkolem MDM je propojení všech kmenových dat, které jsou získány z různých zdrojů, a jejich následné poskytnutí na potřebných místech. MDM zamezuje používání více (nekonzistentních) verzí kmenových dat, které se mohou v organizacích objevit. Organizace se mohou vyhnout situacím, jako je nesprávná komunikace se zákazníkem, chyby v dodávkách zboží a služeb, které způsobily nesprávné údaje či nedostatečná přesnost a kvalita reportování.

Kmenová data (master data)

Kmenová data lze popsat jako data popisující aktiva společnosti. Tyto aktivity mohou být **hmotné** (osoby, kontakty, produkty) či **nehmotné** (finance). Kmenová data lze pro potřeby MDM rozdělit do tří kategorií:

1. **Transakční data** - uchovávají záznamy o aktivitách v business procesech. Primárně vznikají v provozních systémech a typicky nejsou sdíleny s ostatními systémy.
2. **Operativní data** - přiřazují transakčním datům vazby k subjektům (dodavatelům, zákazníkům a dalším). Operativní data jsou sdílena mezi systémy.
3. **Analytická data** - setřídí operativní data do větších celků, které lze řídit a reportovat.

Data lze rozdělit také podle oblastí, které popisují. Těmito oblastem se říká **datové domény**. Příkladem může být zákaznická doména, kam budou patřit data o zákazníkovi. Data se mohou mezi doménami překrývat.

Organizace pracují s daty v různých aplikacích a systémech (např. ERP systémy, databáze, modelovací nástroje, ETL nástroje a další). Tyto data mají různou formu, úroveň detailu a významu, proto jsou tyto informace popsány v **metadatech**. Jelikož jsou metadata implicitně součástí aplikací, MDM vrstva plní roli sdíleného repozitáře metadat.

MDM řešení obsahuje tři dimenze, jsou to:

1. **Master data domény**
2. **Způsoby použití**
3. **Implementační styly**

Ad 1) Master data domény

V rámci vývoje MDM došlo k vytvoření dvou klíčových částí:

- **Zákaznická data - Customer Data Integration (CDI)**
- **Produktová data - Product Information Management (PIM)**

CDI systém se zaměřuje na vedení lidí a organizací. CDI může agregovat data z ostatních systémů, spravovat tato data a předat do dalších systémů (např. účetní systémy, CRM systémy a další). PIM systém spravuje definice a životní cyklus hotových produktů a služeb. PIM systém sbírá informace o produktech z většího počtu zdrojů, dostává schválené definice produktů a ty poté publikuje (např. na webových stránkách, do marketingových systémů atd.).

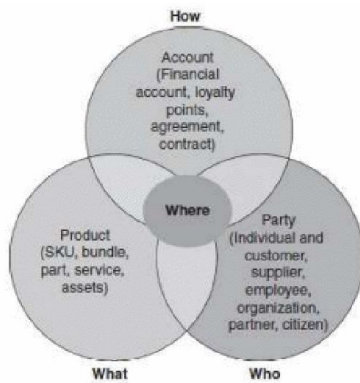
Master data lze rozdělit do tří domén (subjekt, produkt a účet). Všechny domény obsahují určitou třídu informací.

Doména subjekt (Party) - obsahuje informace o osobách nebo organizacích.

Doména produkt - obsahuje informace o produktech a jaké subjekty je nabízejí či používají.

Doména účet - popisuje, jak je subjekt spojen s produktem, který nabízí či vlastní.

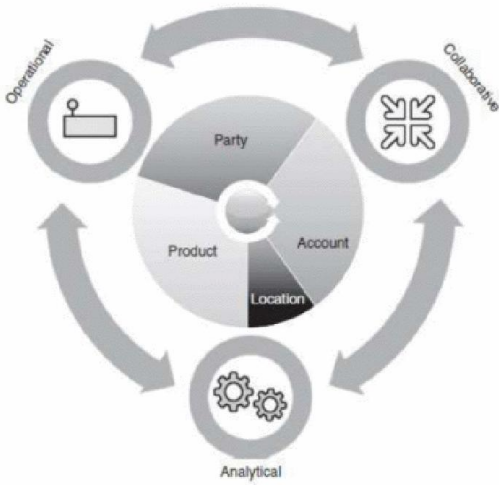
Dále mezi daty mohou být informace o umístění (lokality), které se často spojují s jednou ze tří domén.



Ad 2) Způsoby použití

Jednotlivé formy realizace MDM:

1. **Operativní MDM** - realizováno "proti proudu" obecných datových toků. Forma je reprezentována jedním či více MDM zabudovaným do aplikací pro jejich integraci. Zaměřuje se na poskytnutí jednotného pohledu na kmenová data v jednotlivých systémech. Definice entit se v čase nemění. Příkladem je produktový katalog.
2. **Analytické MDM** - realizováno "po proudu" datových toků. Představuje součást datových skladů. Definice entit je značně proměnná. Příkladem je MDM řešení orientované na data zákazníků.
3. **Kolaborativní MDM** - se zabývá procesy jako autorizace master data. Představuje systém založený na workflow umožňující činnosti jako eskalace kandidátů na sloučení nebo proces schvalování zavedení nového produktu do produktového katalogu.
4. **Enterprise MDM** - představuje autonomní infrastrukturu založenou na integraci dat z různých systémů. Umožňuje obousměrné datové toky. Vytvoří unifikované datové báze směrem po proudu datových toků a propagace informací z MDM do jednotlivých aplikací proti proudu datových toků.



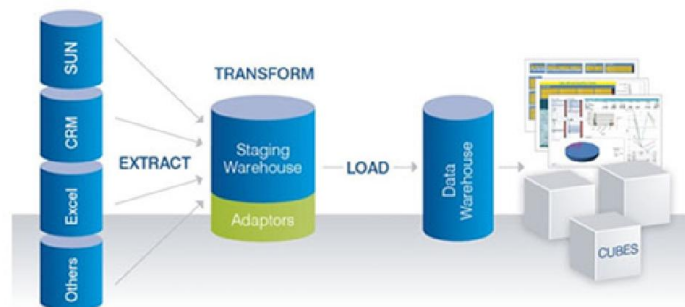
Ad 3) Implementační styly

Podle míry synchronizace metadat s ostatními aplikacemi lze uvažovat o několika implementačních scénářích:

1. **Konsolidace** - je založena na vytváření "*zlatého záznamu*" na základě dat získaných z různých systémů. Změny jsou propagovány po směru datových toků (např. datového skladu), kde proběhne jejich konsolidace a vytvoření unifikované báze.
2. **Forma registru** - poskytuje minimum informací ostatním systémům. Informace je předávána formou reference na jiné aplikace. MDM obsahuje unikátní identifikátor (klíč) k záznamům ve zdrojových systémech. Dotazování probíhá dynamicky ve dvou krocích. Nejprve jsou informace vyhledány v registru a je získán identifikátor. Ve druhém kroku se pomocí identifikátoru získají data přímo ze zdrojových systémů. Poskytnutá sada informací je složena z informací v registru a dodatečných informací ze zdrojových systémů.
3. **Forma koexistence** - MDM řešení je synchronizováno se zdrojovými systémy a předává do nich "*zlatý záznam*". MDM není jediným místem, kde jsou data aktualizována.
4. **Transakční Hub** - Veškerá komunikace prochází přes MDM Hub. Do ostatních systémů jsou propagovány jen jedinečné identifikátory a všechny ostatní atributy jsou poskytovány na základě volání příslušné služby MDM Hub.

ETL

Zkratka ETL je složena ze slov **Extraction, Transformation a Load**, což volně přeloženo znamená získávání, modifikace a uložení. Jedná se o sadu procesů, na jejichž základě jsou data získávána z mnoha různorodých databází, aplikací a systémů, měněna a upravována na základě stanovených pravidel a nakonec uložena do cílových systémů, což nejčastěji bývají datové sklady, datová tržiště, ale stejně tak to mohou být klasické relační databáze či textové soubory.



Extraction

V první fázi ETL procesu je třeba získat potřebná data ze zdrojových systémů. Často se jedná o zcela odlišné zdroje dat, ať už jde o datové sklady, relační databáze, textové soubory v různých formátech, XML soubory, COBOL, ERP, CRM aplikace apod. Každý systém může navíc používat odlišnou organizaci dat a různé datové formáty či být umístěn na rozdílných platformách. V této fázi je klíčové, aby se podařilo získat všechna data ze zdrojových systémů v požadovaném čase, což může být z různých důvodů problematické. Důležité je u každého zdroje definovat dostupnost, kvalitu a obchodní význam jednotlivých položek. Nedílnou součástí extrakce je i ověření, zda data jsou v očekávané struktuře a kvalitě. Případné nedostatky jsou zaznamenány do logů a nejsou zahrnuty do zpracování od počátku. Data, která projdou kontrolami kvality, jsou nahrána do oblasti nazývané obvykle stage. To bývá samostatný databázový stroj, vyhrazený právě pro transformační procesy.

Transformation

Druhou fází jsou transformace a čištění dat. Na začátku je třeba specifikovat obchodní pravidla, která musí definovat sami uživatelé. Tato pravidla říkají, jak zacházet s daty, jejich nekonzistencí, jakým způsobem řešit chybějící data či chyby v nich. V praxi se jedná o pravidla definující referenční integritu, která nám popisuje vazby mezi daty, pravidla pro zacházení s duplicitami či chybějícími daty, dále pravidla pro doménovou integritu, která definují, jakých hodnot či typů hodnot mohou data nabývat, a také pravidla týkající se dat jako takových. Nakonec se musí stanovit pravidla pro mapování zdrojových dat na cílová.

Příklad běžně používaných transformací:

- výpočet nové hodnoty (finální cena = původní cena * (1-sleva)),
- součet dat z více řádků (celkový počet zaměstnanců v pobočce),
- spojení dat z různých zdrojů (jména zaměstnanců, kteří pracují v místě, kde se vyvíjí určitý produkt),
- rozdělení dat do více polí,
- sjednocení formátů a měřítek.

Load

V poslední fázi ETL procesu je potřeba získaná, upravená a pročištěná data uložit do cílového systému (nejčastěji datového skladu). Náročnost tohoto ukládání je různorodá a záleží jak na povaze zdrojových dat, tak na požadavcích na cílový systém. Některé datové sklady pouze přepisují staré informace novými, jindy je třeba dělat inkrementální aktualizace. Existují také více komplexní systémy, které udržují historii a audit všech změn v datech.

Úskalí ETL procesů

Rozdílná terminologie – data mají stejný název, ale obsahují jiné informace	Vícenásobné položky – sloučené adresy
Různé formáty dat – datum, čísla, rodná čísla, PSČ	Nedodržená referenční integrita – porušení vazeb mezi tabulkami
Chybějící data – v primárních systémech chybí potřebná data – například datum	Duplicita dat – nutné určit priority zdrojů
Různá měřítka – nutno převádět na stejné jednotky	Rozdílné kódování
Nejednoznačné údaje – jeden údaj je uložen různými způsoby	

Realizace ETL

Vlastní program/skript

Ve velkém množství projektů datových skladů jsou ETL procesy vyvíjeny a provozovány na bázi skriptů, tj. programů od základů vytvořených vývojovým týmem v některém z obecně používaných jazyků - SQL, C/C++, Perl a podobně. Výhodou tohoto přístupu je minimální prvotní investice, jelikož vývojáři jsou zpravidla schopni vyvíjet skripty bez nutnosti rozsáhlejšího zaškolení a příslušná infrastruktura je již k dispozici.

Velké riziko ovšem tkví v náročné správě a údržbě řešení založeného na skriptech a minimální flexibilitě. Psaní kódu je zdouhavé, náchylné na chyby a klade vysoké nároky na vývojáře. To se promítne v relativně nízké produktivitě a rapidním zvyšování nákladů na ETL řešení.

ETL nástroje

Negativní aspekty vývoje ETL procesů na bázi skriptů přinesly požadavek na zefektivnění tohoto procesu. Několik softwarových firem se úkolu ujalo a uvedly na trh specializované ETL nástroje, které výrazně zvyšují produktivitu vývoje, nabízejí potřebnou flexibilitu a ve výsledku redukuje náklady na vývoj ETL systému.

Použití ETL nástroje přináší hlavně následující výhody:

- **Vysoká produktivita** – přehledné grafické prostředí šetří vývojáři čas,
- **Flexibilita** - Díky objektovému přístupu je velice jednoduché procesy modifikovat, rozšiřovat a přizpůsobovat změnám požadavků a podmínek,
- **Výkon** - ETL nástroje jsou koncipovány tak, aby byly schopny optimálně využít hardwarových a systémových prostředků a dosáhnout maximálního výkonu,
- **Otevřenost** - ETL nástroje zahrnují technologie pro přístup k nejrůznějším typům podnikových systémů,
- **Podpora metadat** –ETL nástroje pracují s popisnými informacemi o zdrojových a cílových objektech, transformačních předpisech, statistikách provozu atd.

Typy ETL nástrojů

Dnes je na trhu několik ETL nástrojů, jejichž společným posláním je ulehčit a zefektivnit vývoj ETL procesů. Liší se rozsahem funkcionality, architekturou, mírou otevřenosti, výkonem a v neposlední řadě také cenou.

1. ETL nástroje **první generace** - na základě graficky navrženého transformačního předpisu vygenerují kód, který se kompiluje a spouští většinou na zdrojové platformě,
2. ETL nástroje **druhé generace** - jádrem je transformační engine realizující ETL procesy podle objektivě definovaného předpisu uloženého v katalogu metadat a graficky navrženého vývojářem.

Zdroje

Systémová integrace

Systémová integrace (System Integration) - <https://managementmania.com/cs/systemova-integrace>

Architektury integrace pro rozsáhlé informační systémy a ekonomické aspekty systémové integrace - <http://www.ekf.vsb.cz/export/sites/ekf/cerei/cs/Papers/VOL13NUM03PAP05.pdf>

Moderní integrace aplikací - http://download.microsoft.com/download/8/6/c/86c09926-affc-4e14-bec0-3c45cd989436/Moderni_integrace.pdf

Replikace - <http://www.cs.vsb.cz/kot/Reports/Dais2/index.htm>

EDI

- <http://www.ccv.cz/>
- <http://www.cvis.cz/hlavni.php/soubory/images/hlavni.php?stranka=novinky/clanek.php&id=512>
- <http://www.systemonline.cz/sprava-dokumentu/elektronicka-vymena-dat-aktualne-a-z-ruznych-pohledu.htm>
- <http://edi.skoda-auto.cz/>
- <http://www.editel.cz/>
- <http://www.makro.cz/pro-dodavatele/edi>
- <http://www.edizone.cz/elektronicka-vymena-dat-edi/edi-ready-retezce/>
- <http://www.edibasics.co.uk/what-is-edi/>
- <http://www.edizone.cz/>
- http://cs.wikipedia.org/wiki/Elektronick%C3%A1_v%C3%BDm%C4%9Bna_dat

MDM

- http://www.unicorncollege.cz/european-ii-center/cesnek-rostislav/attachments/BP_Cesnek_Rostislav.pdf
- http://www.dataquality.cz/index.php?ID=5&clanek=201502_MDM
- <http://www.systemonline.cz/business-intelligence/master-data-management-1.htm>
- http://en.wikipedia.org/wiki/Master_data_management

ETL

- <http://www.systemonline.cz/clanky/co-se-skryva-pod-zkratkou-etl.htm>
- <https://touchstonefms.co.uk/wp-content/uploads/2012/10/etl-agile2.jpg>
- http://www.vse.cz/vskp/show_file.php?soubor_id=1234129
- https://is.bivs.cz/th/6357/bivs_b/SpalenkaOndrej_BIVS_BakalarskaPrace.pdf
- http://is.muni.cz/th/373858/fi_b/tlac_BC.docx