



A MIGRÁCIÓ MÖGÖTT



Napjaink digitális kihívásainak megfelelően egyre több olyan feladattal állunk szemben, amely a tárolt adatok mennyiségének nagyságrendi növekedését hozza magával, miközben az adattartalom is összetettebbé, az adatkörök kapcsolatai is egyre bonyolultabbá válnak. A helyzet hatékony, szervezett és fenntartható kezelése komoly kihívások elé állítja a vállalatokat. Gyakran találják magukat szemben olyan problémákkal, mint pl. a rendszerek fejlesztési korlátai, a felhasznált technológiák elavulása vagy éppen olyan performancia problémák, amelyek kezelése sokszor egy új rendszer bevezetését igényli, ami akár egy teljesen új működési környezet kialakításával is együtt járhat.

Az informatikában az adatmigráció (vagy röviden: migráció) mint fogalom az adatok rendszerszintű áttöltését jelenti. Egy vagy több forrásrendszerből történő adatkinyeréssel, egy vagy több új célrendszer felé történő transzformációs adat áttöltést foglal magába, amely az adatminőség ellenőrzését és/vagy – kontrollált módon – az adatok javítását is szolgálja.

Lássuk, milyen tényezők vezethetnek adatmigrációs helyzetek létrejöttéhez, amelyeknél általános követelményként elmondható, hogy az új rendszernek indulásakor tartalmaznia kell a régi rendszer adatait a rendszer használatához szükséges – akár teljeskörű történeti – adattartalommal:

Üzleti változások:

- A szervezet növekedése, fejlődése performancia problémákat okoz vagy olyan szervezeti változások történnek, amelyeknek megfelelő működés nem képezhető le a rendszerben;
- Piaci igényeknek való megfelelés alapvetően eltérő megközelítést vagy új technológiai és integrációs képességeket igényel;
- A rendszer működtetése, üzemeltetése sok manualitást igényel, nagy és növekvő költségekkel jár;
- Az adat vezérelt működés irányába való elmozdulás koncepcionálisan új adatigényeket és adatkapcsolati leképezéseket igényel;
- Felvásárlás/beolvadás során rendszerkonszolidáció válik szükségessé.

Technikai változások:

- Jelenlegi rendszer(ek) technológiája elavulttá válik, megszűnik a támogatásuk, fenntartásuk növekvő operatív kockázatokat jelent;
- A jelenlegi rendszerek funkcionalitása vagy integrációs képességei nem elegendőek vagy hibásak, és a rendszer továbbfejlesztése magas és egyre növekvő költségekkel jár;
- Belső költségoptimalizációt célzó rendszerkonszolidáció, platform csere, infrastrukturális változások (pl. on-premise <-> Cloud).

Szabályozás megváltozása:

- Az eddig használt rendszerek hazai, vagy EU-s jogszabályok változása által kizárásra kerülnek.

A régi rendszerek új rendszerekkel történő kiváltása minden esetben adatmigrációt igényel, amely során a migrációs folyamatok rugalmas és hatékony informatikai támogatásával jelentősen lehet csökkenteni a kockázatokat és költségeket.

A módszertan

Két technológiai módszertant emelnék ki a migráció megvalósítása tekintetében: a Kinyerés-Átalakítás-Betöltés, vagy angol megfelelőjeként az Extract-Transform-Load (ETL), illetve a Kinyerés-Betöltés-Átalakítás, vagy angol megfelelőjeként Extract-Load-Transform (ELT) módszertanait.

ETL:

A kinyerés, átalakítás és betöltés (ETL) egy olyan adatáramlási megoldás, amely különböző forrásokból származó adatok gyűjtésére szolgál. Az adatok átalakítása üzleti szabályok szerint történik, majd az adatok betöltésre kerülnek egy céladattárba.

Az átalakítási részfolyamat rendszerint egy külön tárhelyen, köztes adatbázisban történik, és emiatt gyakran átmeneti táblákat használ az adatok ideiglenes tárolására, az átalakításra, valamint a célhelyre való betöltés megvalósítására.

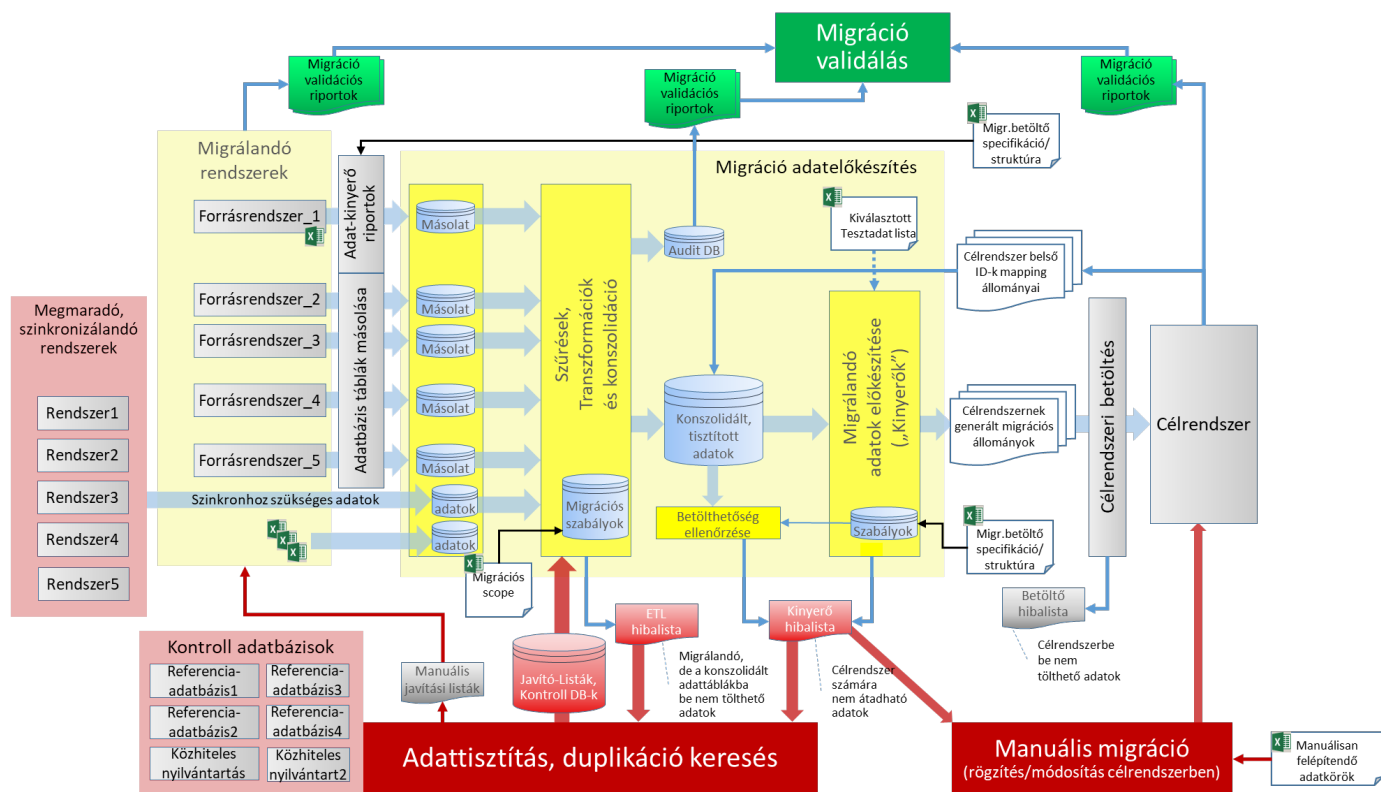
ELT:

Az ELT legfontosabb jellemzője, hogy az átalakítás végrehajtásához használt adattár ugyanaz az adattár, ahol az adatokat majd felhasználják. Az ELT tipikus használati esetei a big data-tartományba esnek. Az összes forrásadat egy skálázható tárolóba kerül áttöltésre a céladattáron belüli külső táblákba, majd ebből a skálázható tárolóból a felhasználásra kialakított céladattár közvetlenül olvassa be az adatokat. Ez a megközelítés kihagyja az ETL-ben található köztes tárolóhelyre történő adatmásolási lépést, amely gyakran időigényes művelet lehet nagy adathalmazok esetében.

A forrásadatok betöltése után a külső táblákban található adatok az adattár képességeivel dolgozhatók fel. Ennek megfelelően kell a céladattár skálázhatóságát megtervezni. Big Data-forgatókönyvekben ez azt jelenti, hogy az adattárnak képesnek kell lennie a nagymértékben párhuzamos feldolgozásra (MPP), amely kisebb adattömbökre bontja az adatokat, és párhuzamosan osztja el az adattömbök feldolgozását több csomópont között.

Az ELT-folyamat utolsó fázisa általában a forrásadatok végleges formátummá alakítása. Előfordulhat például, hogy az adatok particionálva lesznek. Emellett az ELT egyéb optimalizált tárolási formátumokat is használhat. (learn.microsoft.com - Kinyerés, átalakítás és betöltés (ETL), dátum nélk.)

A migrációs módszertannak, illetve a folyamat részeinek és összetettségének könnyebb érthetősége okán mellékelek egy tipikus architektúra ábrát egy ETL típusú migrációs folyamat mentén:



A migrációs folyamat részeinek rövid, tartalmi leírása:

- **Migrálható forrásrendszereknek:** azok az IT rendszerek, amelyek meghatározott adatai a migráció során áttöltésre kerülnek.
- **Megmaradó, szinkronizálható rendszerek:** azok az IT rendszerek, amelyek a migráció szempontjából releváns forrásadattal nem rendelkeznek, de a célrendszer számára szinkron adatokkal szolgálnak (pl. központi partnertár).
- **Kontroll adatbázisok:** olyan adatbázisok, amelyek a migrációhoz kapcsolódó adatellenőrzési és adattisztítási tevékenységek során referencia adatbázisokként szolgálhatnak.
- **Manuális javítási listák:** olyan nyilvántartások, amelyek kezelése nem valamely üzleti rendszerben történik, hanem manuálisan állítják elő, a migráció számára.
- **Migrációs adatelőkészítés:** az a folyamat, amely során forrásrendszerekből kiöntött adatokon megtörténnek a szükséges szűrések, transzformációk és konszolidációk. A migrációs adatelőkészítés végrehajtásáért általában egy migrációs keretrendszer vagy eszköz felelős.
- **Adattisztítás:** egy egész migrációs folyamatot végig kísérő tevékenység. A migrációs projekteknél kiegészítő feladatként szervesen jelen tud lenni az adattisztítás is, melynek szükségességét és mértékét a projekt tervezésekor alaposan fel kell mérni. Az adattisztítási feladatokhoz az inputot egyrészt a felmérés során azonosított, ismert hibák, másrészt a migrációs folyamat során keletkező hibalisták adják.

- **Migrációs hibalisták:** a migrációs folyamat során keletkezett adathibákat meghatározott csoportosítás szerint összesítő, tételes listák.
- **Migráció validálás:** a migráció helyességének ellenőrzése egy független, külső szereplő által.
- **Manuális migráció:** az adatoknak a célrendszer felhasználói felületén keresztül elvégzett adatrögzítését jelent. (Kiss)

Kihívások

Tapasztalataink szerint, az adatmigrációs projektek legnagyobb kihívásai, amelyekre figyelemmel kell lenni az operatív folyamatok kialakítása és a támogató szoftver kiválasztása során a következők:

- **Üzleti területek aktív bevonása:** A migrációs adatmegfeleltetési- és validációs szabályokat az üzleti területek tudják meghatározni. Így érdemes olyan eszközrendszert alkalmazni a migráció támogatására, amelynek használatát az üzleti felhasználók is képesek elsajátítani.
- **Gyakori változások kezelése:** A migrációs és adat validációs szabályok, valamint a migrációs struktúrák jellemzően folyamatosan változnak, finomodnak a migrációs tesztelési ciklusok során. A támogató eszközöknek és a kialakított folyamatoknak képesnek kell lenni a változások rugalmas kezelésére, agilis megközelítési mód támogatására.
- **Adatminőségi problémák kezelése:** A forrás- és célrendszerek adat struktúrái és adatminőséggel kapcsolatos követelményei jellemzően eltérők. Ennél fogva a migrációt támogató eszközöknek és folyamatoknak képesnek kell lenniük a migráció szempontjából kritikus adathibák hatékony feltárására és kezelésére.
- **Komplex migrációs szabályok megvalósítása:** Az egyszerű megfeleltetési logika mellett szinte minden migráció során adódnak olyan migrációs szabályok is, amelyek nagyon komplex megfeleltetési logika implementálását igénylik. Így az alkalmazott eszköznek hatékonyan kell kezelnie az összetett szabályokat, szabályrendszereket is.
- **Megfelelő dokumentáltság biztosítása:** A migrációs teszt ciklusok során a gyakran több száz migrációs állomány és az akár több tízezer mezőmegfeleltetési szabály folyamatosan finomodik, változik. A folyamatot támogató eszköznek biztosítania kell, hogy a migrációs lépések pontosan dokumentáltak legyenek, azok alapján egyértelműen reprodukálható legyen, hogy pontosan mi történt az adott migrációs ciklusban.
- **Kontroll és követés megvalósítása:** A migrációs feladatok hatékony menedzsmentjének fontos előfeltétele az egyes migrációs ciklusok, illetve a teljes migrációs feladatrendszer pontos státuszának és eredményeinek ismerete. A migrációt támogató eszköznek biztosítania kell a folyamat kontrolljához és követéséhez szükséges információk folyamatos rendelkezésre állását.
- **Emberi hibázási lehetőségek minimalizálása:** Egy komplex migrációs folyamat akár több száz lépésből áll, azok bonyolult függőségi hálójával. Az emberi hibázási lehetőségekből adódó problémák minimalizálása érdekében a migrációt támogató eszköznek a folyamat lehető legmagasabb szintű automatizálását kell biztosítania. (Lovász)

A migrációs projekt visszamérése

Migrációs projektekről önmagukban ritkán beszélhetünk, jellemzően valamilyen nagyobb rendszerkiváltási projekt részeként fordulnak elő, de természetesen a migrációs projekteknek is vannak olyan mérőszámai, amelyek a feladat komplexitását, vagy hatékonyságát mutatják a korábban bemutatott migrációs folyamat mentén.

A migrációs projektek hatékonyságát több mérőszám segítségével lehet mérni. Ilyen mérőszámok, nem kizárólagos módon, például a következők lehetnek:

- **Sikeresen migrált rekordok aránya:** Az alkalmazott migrációs megközelítés, illetve migrációs eszköz által sikeresen migrált adatok köre lehetőleg a teljes migrálandó adatmennyiséget fedje le a költséghatékonyság és az átfutási idő kedvező megtartása érdekében.
 - **Hibás, a migrációs eszközzel nem kezelhető adatok száma az összes adathoz viszonyítva:** A mutató az utófeldolgozás, kézi adatrögzítés igényének arányát mutatja, amelyet célszerű a lehető legalacsonyabb szinten tartani a költségek minimalizálása, a manuális rögzítésből adódó hibák elkerülése érdekében.
 - **Kisiklások száma:** Sikertelen migrációs iterációk számát és arányát mutatja, amely az end-to-end migrációs folyamatok és a migrációs logika megvalósításának hatékonyságát mutatja.
 - **Migráció hatása az adatbázis méretére:** Optimalizált formában az újonnan létrehozott adatbázis egyik jellemző mérőszáma lehet.
 - **Migrációból adódó állás idők:** Törekedni kell ennek a mutatónak az alacsonyan tartására, hogy az üzletmenet folytonossága zavartalan maradjon.
 - **Kiegészítő átmeneti tárolóeszközök, hardware-ek, adatbázisok száma:** Törekedni kell ennek a mutatónak az alacsonyan tartására.
 - **Megoldott migrációs hibák száma:** A hibákat rangsorolni és a javításukat optimalizálni kell.
- Tisztított adatok aránya:** Az adattisztítás nagymértékben képes befolyásolni a projekt lefutását, ezért hangsúlyos a helyes arány megválasztása (Levin, 2008) (Fábián)

A megoldás környezete

A sikeresség egyik, és talán legfontosabb kulcsa, hogy a migráció végrehajtásához a támogatást a szervezet több vezetői szintjén is meg kell szerezni. Ehhez elengedhetetlen a megfelelő rendszerességű és az egyes szintek adatigényéhez illeszkedő tartalmú tájékoztatás a szervezeten belül. A migrációs projektek a projekt szponzortól folyamatos és aktív támogatást igényelnek, biztosítani kell számára a folyamat nyomon követéséhez és a szükséges döntések meghozatalához szükséges naprakész információkat.

Ha a vezetői elköteleződés nincs meg, vagy a vezetők nem kaptak megfelelő információkat a döntések meghozatalához, akkor nagyon nagy a projekt sikertelenségének a kockázata. A felsővezetésnek fel kell vállalnia azokat a döntéseket is, amelyek a folytatást, vagy a felfüggesztést eredményezik. Ilyen esetek lehetnek például az alábbiak:

- Adattisztaság elfogadható mértéke
- Ismert hibákból, hiányokból eredő kockázatok felvállalása GO döntésnél
- A teljes migráció ütemezésének jóváhagyása
- Ütemezés, és átfutási idők drasztikus és irreális meghatározása, fejlesztés csúszások kompenzálása

A megoldás eszköze

A migrációs projekteknek jellemzően nagy mind az üzleti-, mind a szervezeti- és folyamat-, valamint a rendszer érintettsége is, ami rendkívül komplex megközelítést igényel mind módszertani, mind az operatív működési és irányítási folyamatok mentén. A Clarity Consulting Kft. egy olyan megoldást kínál, ami az adat migráció folyamatait teljeskörűen lefedi egy kontrollált, automatizált és komplett dokumentációt előállító szoftvertermék segítségével.

A Clarity Consulting saját fejlesztéseként a MIGNON migrációs szoftvercsomag áll rendelkezésre a felsorolt problémakörök hatékony megoldására.

<https://www.datamigrationeasy.com/>

A MIGNON keretrendszer jellemzői:

- Használatát az üzleti felhasználók is képesek elsajátítani
- Képes a változások rugalmas kezelésére, agilis megközelítési mód támogatására
- Képes a migráció szempontjából kritikus adathibák hatékony feltárására és kezelésére
- Hatékonyan kezeli az összetett szabályokat, szabályrendszereket is
- A migrációs lépések pontosan dokumentáltak
- Biztosítja a folyamat kontrolljához és követéséhez szükséges információk folyamatos rendelkezésre állását
- A lehető legmagasabb szintű automatizálását biztosítja

A migrációs feladatrendszert végig kíséri az adattisztítási feladatrendszer. A forrásrendszeri adattisztítási folyamatokhoz célzottan a Clarity DQMPro szoftvere tud segítséget nyújtani.

<https://clarity.hu/szolgaltatas/uzleti-technologiak/adat-es-informacio-menedzsment/>

Ha adatmigrációs témában érintett folyamatokat és lépéseket tervez végrehajtani, és úgy gondolja, hogy a migrációt szakértő kezekre érdemes bízni - elkerülve a tervezés és megvalósítás folyamatában rejlő veszélyeket -, akkor keresse meg a Clarity Consulting Kft-t, és vázoljuk fel együtt a megoldást!

Köszönet Kiss Tamásnak és Lovász Ottónak a cikk létrejöttéhez!

Idézett forrásmunkák

Fábián, T. (dátum nélk.). ADATBÁZIS-MIGRÁCIÓ MÓDSZERTANÁNAK BEMUTATÁSA A DOKSI.HU PÉLDÁJÁN KERESZTÜL. Forrás: doksi.net: <https://doksi.net/hu/get.php?lid=17832>

Kiss, T. (dátum nélk.). személyes tapasztalatok elméleti információk a gyakorlatból megalapozva.

learn.microsoft.com - *Kinyerés, átalakítás és betöltés (ETL)*. (dátum nélk.). Forrás:
<https://learn.microsoft.com/hu-hu/azure/architecture/data-guide/relational-data/etl>

Levin, R. (2008. augusztus 29). *Data migration strategies*:. Forrás: <http://wikibon.org/>:
http://wikibon.org/wiki/v/Data_migration_strategies

Lovász, O. (dátum nélk.). Mignon gyakorlati dokumentáció .