

Vytvorenie Dátového skladu NBS

Prípravné trhové konzultácie



Daniel Martanovič/OIT



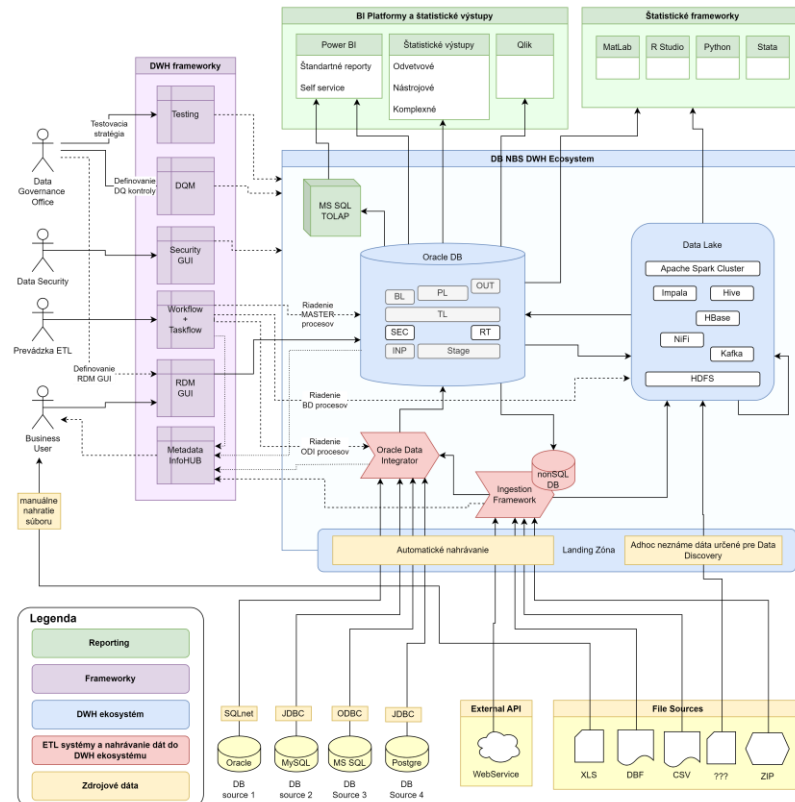
9.5.2022

1. Charakteristika NBS DWH
2. Požiadavky na NBS DWH
 - Aplikačná architektúra (infraštruktúra) NBS DWH
 - Technické požiadavky
 - Business požiadavky
 - Zdrojové systémy
3. Projektový plán a prioritizácia úloh
4. Otázky
5. Diskusia

- Riešenie novej dátovej integrovanej základne (NBS DWH) je navrhnuté ako jeden dátový ekosystém v kombinácii dvoch platforiem:
 - DataLake
 - Relačná časť
- V maximálnej miere sa požaduje použiť výsledky predchádzajúceho projektu „Pilot DWH“:
 - Technická infraštruktúra – Oracle DB, MSSQL/TOLAP a Power BI
 - Know-how pri extrakcii dát z ŠZP a RBUZ
 - ETL nástroj pre NBS DWH bol vybratý Oracle Data Integrator ako výsledok interného výberu

- Predmetom riešenia sú tieto komponenty architektúry:

- **Relačná časť DWH** a jeho dátové vrstvy, dátové toky a metadátové nastavenia
- **Data Lake DWH** a jeho dátové vrstvy, dátové toky a metadátové nastavenia
- **Spoločné frameworky** – aplikácie a nástroje, ktoré nie sú súčasťou samotných platforiem dátovej integrácie, ale využívajú sa pre celkové riadenie spracovania dát, podporu prevádzky a podporu data governance procesov.



- Dodávka technických aplikačných frameworkov
 - Workflow manager
 - Ingestion framework
 - Podpora DQM a automatizácie testov
 - Podpora bezpečnostných pravidiel (viditeľnosť dát)
 - Riadenie dátovej retencie
 - Podpora pre manuálne vstupy a Reference Data Management
 - Metadata management (Používateľský DWH InfoHUB a podpora tvorby obsahu – Glossary, import metadát z ETL a podobne)

- Integrácia zdrojových systémov do jednotného, transparentného a unifikovaného dátového modelu
- Podpora prístupu k dáta cez SQL rozhranie
- Vytvorenie platformy pre Data Discovery (Sand Box)
- Vytvorenie Data Martov pre pravidelný a ad-hoc reporting
- Vytvorenie štandardizovaných exportov z NBS k partnerom NBS
- Podpora pre manuálne vstupovanie dát do DWH
- Transparentnú dokumentáciu DWH nastavenia

Integrované zdrojové systémy

Podľa fáz

Fáza	Počet systémov	Objem dát [GB]	Počet objektov
1	14	929	634
2	15	323	412
3	4	107	100
4	2	50	150
Grand Total	35	1369	1146

Podľa technológie

Technológia	Počet systémov	Objem dát [GB]	Počet objektov
Files	10	811	264
MS Excel	5	16	310
MySQL	4	3	40
Oracle DB	6	351	444
WEB API	9	188	88
Grand Total	35	1369	1146

V rámci dátovej architektúry NBS existujú aj ďalšie databázy a importy z externých zdrojov, ktoré plánujeme počas projektu integrovať. Je ich 28 v rôznom rozsahu náročnosti integrácie. Detaily sú v prílohe „Zdrojové systémy NBS DWH“.

1. S prihliadnutím na požadovanú funkčnosť NBS DWH predpokladáme projekt v trvaní 3 rokov. Je navrhovaný čas postačujúci? Ak nie je, zdôvodnite nám ho prosím.
2. Akú konfiguráciu DataLake odporúčate pri predpoklade 50 zdrojových systémov/1000 objektov/2TB dát mesačne (full extrakt), t.j. max 30TB dát?
3. Identifikovali ste nejakú skutočnosť v dostupných materiáloch k DWH/BI NBS , ktorá by Vám bránila v účasti v plánovanej súťaži alebo túto účasť sťažila? Ak áno, uveďte nám prosím túto skutočnosť, ku ktorej časti v materiáloch prislúcha resp. má súvzťažnosť a tiež prosím, ako ju navrhujete zmeniť, a to tak, aby Vám už nebránila v účasti v plánovanej súťaži alebo aby Vám už nestážovala účasť v plánovanej súťaži.
4. Plánovaná lehota na predloženie ponuky je 60 dní. Je uvedená lehota pre Vás dostatočne dlhá? Ak nie je, uveďte nám prosím dôvod na predĺženie plánovanej lehoty na predkladanie ponúk, a o aký čas ju navrhujete predĺžiť.
5. Aká je podľa Vás indikatívna cena predstaveného riešenia DWH?
6. Aká je podľa Vás indikatívna cena pre servisné služby podpory a údržby plánovaného DWH/BI v rozsahu 8x5 (8 hodín x 5 pracovných dní) na jeden rok?