

Do prípravných trhových konzultácií zákazky na Vytvorenie Dátového skladu NBS sa zapojilo osem hospodárskych subjektov, ktorí zaslali nasledujúce odpovede na zadané otázky:

**Otázka č. 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.**

**Odpovede na otázku č. 1:**

- 1.1 "Na základe našich skúseností považujeme navrhovaný čas postačujúci pre samotnú realizáciu dodávky. Odporúčame však navrhovaný projektový plán rozšíriť, a to nasledovne:
  - Analýza prostredia, overenie predpokladov a 14 zdrojových systémov – 3 mesiace
  - Overenie predpokladov o zabezpečení a prístupoch a nastavenie zabezpečenia riešenia – 1 mesiac
  - Školenia pre používateľov – 2 týždne / fáza (spolu 2 mesiace)
  - Príprava plánu ďalšieho rozvoja riešenia – 1 mesiac
  - Zaškolenie a tréning obslužných pracovníkov pre frameworky – 1 mesiac
  - Projektová rezerva – 1 mesiac
  - Návrh a vytvorenie Data Lake – 3 mesiace".
- 1.2 Navrhované trvanie 3 roky je na základe našich skúseností postačujúce. Veľmi pravdepodobne však bude vyžadovať paralelizmus medzi fázami (napr. paralelný beh UAT I. fázy a analýzy 2. fázy). Ako hlavné riziko dodržania termínov vnímame dostupnosť kvalifikovaných špecialistov v správnom čase a množstve (týka sa aj tímu dodávateľa a aj tímu NBS).
- 1.3 Vzhľadom na požadovanú funkčnosť a rozsah NBS/DWH si myslíme, že navrhovaný čas je postačujúci. Aj vzhľadom na možnosť vývoja niektorých častí paralelne.
- 1.4 V princípe je vybudovanie DWH a Data Lake v požadovaných intenciách realizovateľné. Celkový plán projektu môže byť ale ovplyvnený aj súčinnosťou banky, dostupnosťou rozhraní pre integráciu najmä externých zdrojov, rozsahom integrácie externých zdrojov. V rámci výberového konania tiež očakávame formuláciu detailnejších požiadaviek – viď odpoveď k otázke č. 3.
- 1.5 Navrhovaný čas hodnotíme ako veľmi ambiciózny. Projekt zavádza množstvo nových systémov, čo si vyžiada výraznú súčinnosť NBS v projekte, adopciu používateľov nových systémov a zmeny pracovných postupov.
- 1.6 Áno, navrhovaný čas je postačujúci.
- 1.7 Zo skúseností s budovaním obdobne veľkých dátových skladov je trojročné trvanie projektu dosiahnuteľné, avšak vyžaduje perfektné projektové riadenie a zo strany NBS významnú a efektívnu súčinnosť. Zásadnú úlohu zohráva tiež skutočnosť, či termín ukončenia bude dosiahnuteľný v časovom súlade s procesmi a postupmi NBS (napr. akceptácia, príprava rozhrania zo strany poskytovateľov zdrojových systémov).
- 1.8 Po preštudovaní poskytnutého dokumentu s popisom Dátového skladu DWH/BI NBS je podľa skúsenosti IBM s podobnými projektami rámcové trvanie 3och rokov pre implementáciu riešenia zvyčajne postačujúce. Po detailnej analýze požiadaviek môže však byť potvrdené

trvanie kratšie alebo (vo výnimočných prípadoch) dlhšie. Pre vylúčenie pochybností v uvedenom časovom rámci nie sú servisné služby podpory a údržby.

**Otázka č. 2: Akú konfiguráciu DataLake odporúčate pri predpoklade 50 zdrojových systémov/1000 objektov/2TB dát mesačne (full extrakt), t.j. max 30TB dát?**

**Odpovede na otázku č. 2:**

2.1 Pre realizáciu Enterprise Data Lake On-Premise riešenia odporúčame použiť balík Cloudera Data Platform Private Cloud (CDP). Táto platforma poskytuje všetky potrebné nástroje pre vytvorenie, správu a prevádzku Data Lake platformy.

Všeobecné odporúčania

Všeobecne sa odporúča vytvárať prostredia s trojitou redundanciou dát uložených v HDFS a 30% rezervou pre výpočty.

Dosiahnutie vysokej dostupnosti prostredí Data Lake umiestnením serverov klastra do dvoch rôznych lokalít (stretch cluster) je možné v prípade, že budú sieťové parametre dostatočné pre prenos potrebného objemu dát nahrávaných do prostredia (sieťové spojenie medzi lokalitami by malo mať podobnú prenosovú rýchlosť a kapacitu ako dátové prepojenie rackov v dátovom centre).

Predpoklady použité pri návrhu konfigurácie

Predpokladáme, že jednotlivé prostredia (PROD, TEST, DEV) budú mať nasledovné funkcie:

- Produkčné prostredie (PROD) obsahuje produkčné dáta, ktoré sa načítavajú, transformujú a udržiavajú väčšinou automatizovanými procesmi. Produkčné prostredie podporuje aj produkčné dátové služby, ako sú reporting, dotazovanie, modelovanie alebo export dát.

Vstup do prípravných trhových konzultácií pre účel dodania centrálného celobankového dátového skladu (NBS DWH) - Národná banka Slovenska

Súčasťou produkčného prostredia je aj SandBox, ktorý obsahuje údaje z PROD prostredia. Z dôvodu transformácií a kontrol a kopírovania údajov do SandBox prostredí predpokladáme, že požadovaný objem dát (30 TB) bude v prostredí uložený 2x (s trojitou fyzickou replikáciou a 30% rezervy), t.j. celkovo 240 TB RAW data.

- Testovacie prostredie (TEST) sa používa na overenie nových alebo zmenených procesov nahrávania, transformácie, spracovania alebo modelovania údajov pred ich nasadením na produkciu. Testovacie prostredie bude použité aj na performance testy s produkčnými objemom dát. Predpokladáme, že konfigurácia testovacieho prostredia bude cca 25% kapacity produkcie, t.j. 60 TB RAW data.

- Vývojové prostredie (DEV) sa používa na prototypovanie, vývoj a jednotkové testovanie nahrávania, transformácie, spracovania alebo modelovania údajov. Vývojové prostredie má len obmedzené zdroje a preto nie je vhodné na testovanie vo väčšom rozsahu. Prostredie nebude musieť spĺňať požiadavky na vysokú dostupnosť a konfigurácia nebude vhodná na performance testy. Predpokladáme, že kapacita diskov vývojového prostredia bude cca 10TB RAW data.

Navrhované parametre prostredí

Na základe našich skúseností a informácií poskytovaných spoločnosťou Cloudera sme zostavili odporúčania pre konfiguráciu jednotlivých prostredí.

Pri návrhu Hadoop platformy je potrebné okrem zamýšľaného účelu použitia a plánovaného objemu spracovávaných dát zohľadniť aj detailné požiadavky jednotlivých Hadoop komponentov.

Z týchto požiadaviek vyplýva potreba rozdeliť očakávanú IO záťaž Hadoop komponentov medzi viaceré disky serverov, pričom pri niektorých komponentoch je dôležité mať na IO zdroje výhradný, resp. prioritný prístup. Dedikovaním diskov vieme dosiahnuť prioritizáciu IO výkonu pre komponenty profitujúce z takto alokovaných zdrojov, čo prispeje k zvýšenej výkonnosti, k lepšej predikovateľnosti a stabilite behu Hadoop komponentov.

Rovnaký princíp prioritizácie potrieb vybraných komponentov zohľadnujeme pri návrhu na počet a spôsob záťaže CPU.

V prípade RAM očakávame, že prostredia pobežia v minimálnych konfiguráciách; pri RAM teda počítame do istej rozumnej miery, s možnosťou zdieľania medzi komponentami (tzv. memory overcommit princíp).

Pri zvyšovaní záťaže prostredia predpokladáme, že RAM bude zdroj, ktorý bude treba navýšiť najskôr.

Navrhované parametre prostredí

Produkčné prostredie (parametre sú pre každý zo serverov – nodes): 3x MasterNodes – 32 VCPU, 64GB RAM, 5 diskov (každý s kapacitou 256 GB) 4x WorkerNodes – 32 VCPU, 64GB RAM, 10 diskov (každý s kapacitou 6 TB) 2x EdgeNode – 4 VCPU, 32GB RAM, 1,5 TB diskového priestoru

TEST (parametre sú pre každý zo serverov – nodes): 1x MasterNode – 32CPU, 64GB RAM, 5 diskov (každý s kapacitou 256 GB)

Vstup do prípravných trhových konzultácií pre účel dodania centrálného celobankového dátového

skladu (NBS DWH) - Národná banka Slovenska

2x MasterNodes – 32 VCPU, 64GB RAM, 10 diskov (každý s kapacitou 2 TB) 1x WorkerNode – 32 VCPU, 64GB RAM, 10 diskov (každý s kapacitou 2 TB) DEV (parametre sú pre každý zo serverov – nodes): 3x MasterNodes + Worker Nodes – 32 VCPU, 64GB RAM, 10 diskov (každý s kapacitou 400 GB).

2.2 Pre Datalake odporúčame využiť produkt Cludera Data Platform Private Cloud. Produkt spiria po Zachovaní vlastnosti a funkcionality.

2.3 Konfigurácia DataLake by sme odporúčali 2 alternatívy. Bud' optimálne, alebo komfortné:

Optimálne riešenie: HW komponent	Typ	Parametre
DB server	Výrobca/ verzia	Oracle 19c EE + Partitioning option
Operačný systém		Oracle Linux 8 (RedHat)
RAM (Typ, počet)		64GB
CPU (Typ, počet)		1xCPU Xeon Gold 6244 (8 cores), 3,6 GHz; 1 slot ešte voľný
Súčasná utilizácia CPU		v maxime 30%
Veľkosť DB		2TB (dáta a indexy), 30 TB na rok vzhľadom na úroveň agregácie
Diskové pole (Typ HDD, počet, veľkosť)		28 x 3 TB SSD lokálne disky, zapojené do RAID6; cez PCI RAID radič, je možné zriadiť prístup aj na centrálné diskové pole 3PAR, lebo server má 2x 16 Gbit FC karty
LAN na diskové pole (rýchlosť, počet)		-

LAN do siete (rýchlosť, počet)	4x 10GBit	
<b>BI server</b>	Výrobca/ verzia	MS SQL Server 2019 Enterprise
Operačný systém	Windows Server 2022 Data Center edícia.	
RAM	32 GB	
CPU	4 vCPU	
Diskové pole	3PAR	
<b>Aplikačný server</b>	Výrobca/ verzia	Apache Tomcat/9.0.56
Operačný systém	Red Hat Enterprise Linux 7.9	
RAM	10 GB	
CPU	6 vCPU, Intel(R) Xeon(R) Gold 6134 CPU @ 3.20GHz	
HDD	25 GB	

Komfortné riešenie: HW komponent	Typ	Parametre
<b>DB server</b>	Výrobca/ verzia	Oracle 19c EE + Partitioning option
Operačný systém	Oracle Linux 8 (RedHat)	
RAM (Typ, počet)	128 GB	
CPU (Typ, počet)	2xCPU Xeon Gold 6244 (8 cores), 3,6 GHz;	
Súčasná utilizácia CPU	v maxime 30%	
Veľkosť DB	2TB (dáta a indexy) 30 TB na rok vzhľadom na úroveň agregácie	
Diskové pole (Typ HDD, počet, veľkosť)	28 x 3 TB SSD lokálne disky, zapojené do RAID6, cez PCI RAID radič, je možné zriadiť prístup aj na centrálné diskové pole 3PAR, lebo server má 2x 16 Gbit FC karty	
LAN na diskové pole (rýchlosť, počet)	-	
LAN do siete (rýchlosť, počet)	4x 10GBit	
<b>ETL server</b>	Výrobca/ verzia	MS SQL Server 2019 Enterprise
Operačný systém	Windows Server 2022 Data Center edícia.	
RAM	32 GB	
CPU	4 vCPU	
Diskové pole	3PAR	
<b>Aplikačný server</b>	Výrobca/ verzia	Apache Tomcat/9.0.56
Operačný systém	Red Hat Enterprise Linux 7.9	
RAM	10 GB	
CPU	6 vCPU, Intel(R) Xeon(R) Gold 6134 CPU @ 3.20GHz	
HDD	25 B	

2.4 Nie sme schopní za poskytnutý čas zodpovedne vypracovať odporúčanie.

2.5 Pre Datalake odporúčame využiť produkt Cloudera Data Platform Private Cloud. Produkt spĺňa požadované vlastnosti a funkcionality.

Na základe našich skúseností ako aj na základe odporúčaní spoločnosti Cloudera navrhujeme nasledovnú konfiguráciu:

Produkčné prostredie (parametre sú pre každý zo serverov - nodes):

3x MasterNodes - 32CPU,64GB RAM, 5 diskov,  
4x WorkerNodes - 32CPU,64GB RAM, 10 diskov,  
2x EdgeNode - 4 CPU, 32GB RAM, 2 disky.

TEST (parametre sú pre každý zo serverov - nodes):

1x MasterNode - 32CPU,64GB RAM, 5 diskov,  
2x MasterNodes - 32CPU,64GB RAM, 10 diskov,  
1x WorkerNode - 32CPU,64GB RAM, 10 diskov.

DEV (parametre sú pre každý zo serverov - nodes):

3x MasterNodes \* Worker Nodes - 32CPU, 64GB RAM, 10 diskov.

2.6 Predpokladáme, že otázka sa týka primárne hardvér konfigurácie navrhnutých klastrov a že riešenie by mohlo smerovať k dodávateľovi Cloudera. V dokumente „Príloha č. 1 – Dátový sklad DWH/BI NBS“ je popísané rozhodnutie vytvorenia DataLake klastra „rozkročeného“ medzi dve lokality a prevádzkovanom na virtuálnom prostredí.

Podľa spôsobu využitia uložených dát a hardvéru možností prevádzkovateľa odporúčame upraviť navrhnutú konfiguráciu navýšením diskovej kapacity na min. 130 TB (vysvetlenie nižšie), prípadne prestavať cluster na 6 (dáta) nodový s benefitom využitia Erasure Coding (opäť popísané nižšie).

Počet workerov: Zvolená konfigurácia produkčného klastra so 4 dátami nody (tzv. worker) je minimálna možná konfigurácia ktorá pri nastavení replikačného faktora 3 ochraňuje dáta pred stratou v prípade poruchy jedného z dátových diskov klastra. Storage: U takto malých (čo do počtu workerov) klastrov je nevyhnutné počítať dostatočne veľkú rezervu na diskovom poli, ktorá je, okrem odporúčaných 20 % kapacity celého klastra, ešte navýšená o rezervu vo veľkosti 1 / počet serverov kapacity celého klastra. Podľa očakávaného charakteru práce s dátami by sme ďalej navrhovali navýšenie počtu workerov na 6 tak, aby bolo možné nahradiť replikačný faktor (3) inou data protection metódou, a to metódou Erasure Coding (3,2), ktorá umožní podľa charakteru dát uložiť na rovnaký diskový priestor takmer dvojnásobok dát. Na základe nasledujúcich vstupov ako: formát uloženia dát (parquet, avro, ORC), charakter dát a použitie kompresie (napr. Snappy), kompakcie jednotlivých dátových blokov do partitions a rýchlosti zaplnenia klastra dátami (obvykle nie je v prvom roku potreba „plnej“ kapacity klastra), je možné následne vykonať finálne odporúčanie o veľkosti úložiska. Ak vezmeme do úvahy iba zadanie bez väčšieho kontextu, tak 30 TB business dát po replikácii (3) bude znamenať 90 TB, a pripočítajme rezervu pre prípad pádu jedného nodu klastra (20 TB). Tým sa dostávame na 110 TB, plus 20% operatívna rezerva sa jedná cez 130 TB diskovej kapacity na všetkých dátových nodoch. Z uvedeného zjednodušeného aspektu disková kapacita produkčného servera je nedostatočná. Odporúčame taktiež vykonať rozdelenie dátových diskov (v podstate sa môže jednať o platňové disky) a diskov pre OS (SSD, prípadne SAS disky). CPU: Sizing 8 cores je dostatočný pre beh väčšiny štandardných aplikácií big data klastra a základnú dátovú prácu. RAM – Master node: Najnáročnejší SW komponent je HDFS Name Node, ktorý za bežnej prevádzky potrebuje až 8 GB RAM pre svoj beh, minimálne dokáže fungovať pri cca 1 GB RAM (okrem iného v závislosti od množstva dátumu blokov na HDFS). Minimálna konfigurácia master nodu so všetkými základnými službami je približne 10 GB RAM, optimum môže byť približne na trojnásobku. Z toho dôvodu veríme, že 64 GB RAM bude viac ako dostatočné a mohlo by aj

dôjsť k zníženiu. RAM – Worker node: 64 GB by malo byť dostatočné na väčšinu spracovávaných úloh. Tu by sme odporúčali vykonať nastavenie YARN resource poolu, pretože Spark obvykle alokuje všetku dostupnú pamäť.

Pri testovaní, a tým skôr potom aj vývojovom klastrí, máme vážne pochybnosti, že klaster bude schopný prevádzkovať všetok softvér, ktorý bude nasadený na produkčnom klastrí, a tým pádom nebude dostatočne slúžiť na otestovanie nového vývoja. Častou praxou, vedúcou k zníženiu HW nárokov na neprodukčné prostredie, je zníženie replikačného faktora na 2 (1), a tým ušetriť diskovú kapacitu, ktorá pri stavbe klastra môže byť vo virtuálnom prostredí najdrahšia (pokial' ide napríklad o rýchle SAN pole).

Vnímame túto oblasť ako komplexnejšiu a môžeme v prípade potreby ponúknuť ďalšie konzultácie napríklad formou krátko workshopu.

2.7 Vzhľadom na rozsiahlosť požiadaviek a generického popisu Data Lake je potrebnějšía podrobnejšia analýza a detailnejšie informácie týkajúce sa napr. zdrojových systémov, využitia objektového úložiska a pod. Z popisu jasne nevyplýva, akú technologickú platformu má DataLake využívať. Ak uvažujeme DataLake v kombinácii napr. na cloudovej platforme AWS S3 + spark pre processing a prípravu konsolidovaných dát do DWH, tak rozhodujúcim kritériom nie je ani tak samotná konfigurácia DataLake, ako licenčný model poskytovateľa cloudovej služby.

## 2.8 Konfigurácia v zadaní vo svojej podstate odpovedá predloženým požiadavkám, tj.:

- vybudovanie Data Lake platformy na báze Enterprise balíka (nie čistý Open Source),
- prevádzkovanie 3 oddelených prostredí,
- vysoká dostupnosť s využitím sekundárnej lokácie.

Konkrétna konfigurácia by mala vychádzať z odporúčaní vendora danej distribúcie – napr. Cloudera - odporúča sa použitie serverov strednej triedy.

K návrhu konfigurácie máme tieto odporúčania:

- Vystavanie BigData platformy má zmysel riešiť v 2 fyzicky oddelených lokalitách ak ma byť platforma v postavení business kritického systému. (V prípade sekundárnej lokality je nutné zabezpečiť prepojenie medzi lokalitami aspoň 10 GbE s spojením s vlastnou sieťou)
- Pre produkčné prostredie nie je odporúčané vystavať na virtualizačných platformách. V prípade BigData technológií je odporúčané použiť fyzické servery pre ich predvídateľný výkon.
- Testovacie prostredie odporúčame vystavať na rovnakej platforme ako produkčné (fyzické servery). Najmä z dôvodov aby chyby vzniknuté na produkcii spôsobené HW boli replikovateľné na testovacom prostredí (vylúčiť pri hľadaní možných problémov HW špecifikáciu). Toto je len odporúčanie.
- Sizing pre Produkčné prostredie odporúčame (pre worker nody datalake platformy odporúčame minimálne **256GB RAM** a **16x CPU**).

Prostredie	Uzol	Počet	CPU	RAM	HDD / uzol	Komentár
Produkcia	Worker Node	4	20 cores (fyzické servery)	256 GB	10x 2.4TB HDD 10K + 480GB SSD	Konfigurácia BD platformy pre predpokladanú

	Master Node	2	8 cores (fyzické servery)	32 GB	2x 480GB SSD + 480GB OS	veľkosť dát NBS s funkciou vysokej dostupnosti na serveroch strednej triedy (napr. Dell EMC PowerEdge R550 Rack Server) PROD HW špecifiká: Sieťová karta : 2x10Gb; Worker Node = 2U s možnosťou min 16 slotov na disk; Edge, UI, Master = 1U stačí do 4 diskov Ostatné prostredia: Sieťová karta : 2x10Gb; 1U stačí do 4 diskov
	Edge Node / UI Node	2	8 cores (fyzicky)	64 GB	3x 1.2TB HDD 10K + 480GB SSD	
Dev	UI/Edge/Worker /Master Node	3	8 cores (môže byť virtuál)	32 GB	3x 1.2TB HDD 10K + 480 GB SSD	minimalistická konfigurácia pre vývoj; services rozložené cez 3 uzly
Test	Worker Node	3	8 cores (fyzické servery)	32 GB	3x 1.2TB HDD 10K + 480 GB SSD	zjednotenie PROD a TEST v konfigurácii (potreba minimálnej HA aj v teste)
	UI/Edge /Master Nod	2	8 cores (fyzické servery)	32 GB	2x 1.2TB HDD 10K + 480 GB SSD	8 cores (fyzické servery)

**Otázka č. 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 nesťažovala účasť v plánovanej súťaži.**

**Odpovede na otázku č. 3:**

3.1 Neidentifikovali sme v dostupných materiáloch k DWH/BI NBS žiadnu relevantnú skutočnosť, ktorá by nám bránila v účasti v plánovanej súťaži alebo túto účasť sťažila. Identifikovali sme potenciálne úzke hrdlo v hardvérovej konfigurácii riešenia, nakoľko navrhované hardvérové riešenie pre DWH časť sítě potenciálne postačuje pre popisované riešenie, avšak

pravdepodobne neprináša dostatočnú rezervu pre zabezpečenie pružných reakcií systému na dotazy väčšieho počtu používateľov, ako aj na ďalšiu rozšíriteľnosť a udržateľnosť riešenia v budúcnosti. Jedným z možných alternatívnych riešení je i platforma Oracle Exadata, ktorá umožňuje využiť inžinierované riešenie prinášajúce maximálnu kompatibilitu a výkon pre časť DWH. Hardvérovú konfiguráciu navrhnutú pre Data Lake sme pripomenovali vyššie.

3.2 V dostupných materiáloch sme neidentifikovali žiadnu skutočnosť, ktorá by brzdila našej účasti v plánovanej súťaži alebo ju sťažila.

3.3 Nezistili sme skutočnosti, ktoré by nám bránili v plánovanej súťaži. Prípadné konkrétne detaily by boli riešené v rámci konzultácií.

3.4 Nevidíme prekážku v účasti na plánovanej súťaži.

Pre samotnú súťaž si však dovoľujeme odporučiť doplnenie niektorých detailov, vďaka ktorým bude možné zostaviť ponuku presnejšie a s menším rizikom nevhodne definovaného rozsahu plnenia a celkovo projektového plánu:

Časť „Výstupné dátové oblasti“

- Nie je jasný funkčný rozsah v DWH - či má vzniknúť aplikácia pre správu mapovaní, spájání, agregáciu dát, matematické operácie, formátovanie a publikovanie do výstupných formátov (toto sa zvyčajne robí špeciálnymi aplikáciami), alebo DWH bude implementovať potrebnú prípravu výstupov s fixne danými transformáciami dát.
- či sa požaduje logika konverzie výstupov DWH do formátov GESMES/TS, SDMX/XML, alebo je táto funkcionálna už implementovaná v inom nástroji banky

Rozsah implementácie

- Najmä pre externé dátové zdroje odporúčame špecifikovať rozsah požadovanej integrácie (napr. konkrétne rozhranie, špecifikácia data setov)
- Data marty – bližší detail k DM\_SUP\_\_ESMA

Vo všeobecnosti predpokladáme, že v rámci výberového konania budú dodané prípadné ďalšie detaily a požiadavky s možnosťou dodatočných konzultácií.

3.5 Výrazná časť práce na spracovaní ponuky sa bude odvíjať od merateľných ukazovateľov poskytnutých zadávateľom. Z tohto pohľadu navrhujeme:

Zafixovať množinu integrovaných zdrojových systémov.

Doplniť predpokladané počty objektov jednotlivých zdrojových systémov.

Poskytnúť informácie o odhadovanom počte objektov v jednotlivých vrstvách riešenia. Z popisu nie je jasné či už prebehla fáza analýzy primárneho zdroja pre MDM a prepojenia číselníkov rôznych zdrojových systémov, alebo má byť súčasťou dodávky.

Pre potreby ponuky navrhujeme vymedziť, aké licencie sú jej súčasťou.

3.6 Áno. Výber konkrétnej technológie výrazne limituje uchádzačov (napr. Oracle DB resp. Oracle Data Integrator). Navrhujeme umožniť vlastný výber nástrojov a technických komponentov a zároveň umožniť vybudovanie prostredia vo verejnom cloude s požadovanou bezpečnosťou.



3.7 Podkladová dokumentácia je z nášho pohľadu prehľadne spracovaná. V tejto chvíli nevidíme žiadny problematický bod, ktorý by nám znemožňoval účasť v súťaži. Z hľadiska dopytovaného riešenia veríme, že máme dlhoročné skúsenosti s budovaním obdobných systémov a z prostredia Národnej banky Slovenska. Do istej miery vnímame zvýšené riziká spojené so značným rozsahom dopytu a odporúčili by sme zákazku rozdeliť do čiastkových zákaziek (napr. podľa definovaných fáz). Kľúčovým faktorom pre ne/účasť v súťaži môžu byť kvalifikačné kritériá, ktoré ale zatiaľ nie sú súčasťou dostupných materiálov a budeme ich tak môcť preveriť až v rámci výberového konania.

3.8 V dostupnej dokumentácii pre popis je sú z nášho pohľadu dostatočne špecifikované nasledovné oblasti:

- detailnejší popis "Fázy 0" analyzujúci aktuálnu situáciu aj s verifikáciou výberu vhodných systémov pre úvodnú fázu implementácie (napr. s ohľadom na vytváranie core dimenzií pre relačnú časť),
- detailnejší popis týkajúci sa aktualizácií dát s ohľadom na množstvo dát v dennej delte a časových okien pre spracovanie (dostupnosť aplikácií/systémov),
- chýbajúce akceptačné kritériá,
- "chýbajúce Ne-funkčné požiadavky (Non Functional Requirements) a teda požiadavky:
  - na zálohovanie, odozvy riešenia, resp. odozvu systému v jednotlivých vrstvách - od transferu dát až po reporting,
  - detailnejšia špecifikácia požiadaviek na zálohovanie a záťaž resp. performance systému,
  - požiadavky na servisné služby podpory a pod.,
  - detailnejšie podmienky a parametre migrácie (napr. rozsah iniciálneho loadu)".

**Otázka č. 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.**

**Odpovede na otázku č. 4:**

4.1 Uvedenú lehotu 60 dní na predloženie ponuky považujeme vo všeobecnosti za dostatočne dlhú. V prípade, že by lehota mala začať v období zimných sviatkov alebo letných prázdnin, pre plánované dovolenky by sme odporúčali jej predĺženie o obdobie sviatkov, respektíve o približne 10 dní letných dovoleniek.

4.2 Lehota 60 dní je dostatočne dlhá. Ak by sa malo jednáť o mesiace júl a august, tak navrhujeme pridať extra 3 týždne v septembri. Dôvodom je sada plánovaných dovoleniek a iných projektových aktivít.

4.3 Plánovaná lehota je postačujúca na predloženie ponúk.

4.4 Danú lehotu považujeme za dostatočnú.

4.5 Uvedená lehota je dostatočná.

4.6 Lehota na predloženie ponuky 60 dní je postačujúca, no v prípade ak by bolo obstarávanie v čase dovoleniek, treba zvážiť predĺženie termínu na 80 dní.

4.7 "Podľa našich skúseností s výberovými konaniami pre štátnu správu predpokladáme, že požadovaná ponuka sa bude skladať primárne z ponúkanej ceny rozpadnutej do požadovaného detailu určeného zo strany NBS a naplnenia kvalifikačných kritérií. Z tohto pohľadu sa nám javí 60 dní ako dostatočne dlhá doba. Ak budú na ponuku požadované ďalšie prílohy/dokumenty, môže to mať vplyv na potrebný čas spracovania. Avšak aj v tomto prípade sme pripravení vyvinúť maximálne úsilie, aby sme plánovanú dobu dodržali."

4.8 Áno. Kedy predkladá NBS vyhlásenie verejného obstarávania a akou formou?

**Otázka č. 5: Aká je podľa Vás indikatívna cena predstaveného riešenia DWH?**

**Odpovede na otázku č.5:**

5.1 Cenu predstaveného riešenia DWH odhadujeme v rozsahu 6,4 – 7,2 miliónov EUR (bez DPH).

5.2 Indikatívnu cenu predstaveného riešenia odhadujeme na 5,0 až 5,5 mil EUR bez DPH. Indikatívna cena pokrýva implementačné práce, bez nákupu HW a SW licencií.

5.3 Vzhľadom na veľkosť a rozsah prác sa indikatívna cena veľmi ťažko určuje, no predpokladáme ju cca. 4,5 – 6 MIL EUR bez DPH za dielo.

5.4 Cena predstaveného riešenia, vrátane všetkých súvisiacich služieb, licencií a obstaraní hardware je uvedená v nasledujúcej tabuľke.

<b>Položka</b>	<b>Jednorázová cena bez DPH</b>	<b>Ročná cena / ročný poplatok bez DPH</b>
Obsataranie HW	130.000 EUR	
Prevádzka HW		18.000 EUR
SW licencie jednorázové	500.000 EUR	
SW licencie ročné, poplatky		600.000 EUR
Implementačné náklady	6.600.0 R	

5.5 Hrubý cenový odhad riešenia je viac ako 7.000.000,- EUR bez DPH.

5.6 Indikatívnu cenu riešenia DWH odhadujeme 5-6 mil EUR bez DPH. Cena neobsahuje HW a SW infraštruktúru.

**Otázka č. 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?**

**Odpovede na otázku č. 6:**

6.1 Indikatívnu cenu pre servisné služby podpory a údržby plánovaného DWH/BI v rozsahu 8x5 na jeden rok odhadujeme na 500 - 700 tisíc EUR (ročne).

6.2 Indikatívnu cenu pre servisné služby podpory a údržby plánovaného DWH/BI v rozsahu 8x5 na jeden rok odhadujeme približne na 10% z ceny riešenia. Presnejší odhad bude možné

spraviť po stanovení SLA, komerčných podmienok ako aj odhadovanom množstve práce na rozvojových aktivitách.

- 6.3 Indikatívnu cenu pre servisné služby na 1 rok predpokladáme cca. 150 až 200. 000 EUR bez DPH. Všetky uvedené sumy sú v cenách roku 2022 a nezohľadňujú mieru inflácie, t.j. ak by sa projekt realizoval neskôr je potrebné kalkulovať s navýšovaním cien.
- 6.4 Indikatívna cena pre servisné služby v požadovanom rozsahu 8x5 ročne je cca 27.000 EUR bez DPH. Presné stanovenie ceny môže kolísať na základe požadovaných reakčných časov v definícii podpory.
- 6.5 Hrubý cenový odhad pre servisné služby je 15.000,- EUR bez DPH mesačne.
- 6.6 Štandardne sa pohybuje ročná cena podpory a údržby vo Vami požadovanom rozsahu a SLA 10% z ceny diela.