

**FUNKCIONALNE SPECIFIKACIJE ZA IZVEDBO DRUGE FAZE  
IMPLEMENTACIJE BIGDATA PLATFORME**

Ljubljana, 22. 10. 2025

## 1. REVIZIJA DOKUMENTA

Revizija	Datum	Prispeval	Vsebina
1.0	21.10.2025	Janez Bartol	Prvi osnutek dokumenta
1.1	18.11.2025	Janez Bartol	Dopolnjene specifikacije
2.0	8.1.2026	Janez Bartol	Dopolnjene specifikacije po komentarjih
2.1	19.1.2026	Janez Bartol	Pregled in urejanje

## Vsebina

1.	REVIZIJA DOKUMENTA .....	2
2.	UVOD .....	4
3.	REZULTATI PILOTNEGA PROJEKTA.....	5
3.1.	Vzpostavljeni dostopi do virov podatkov .....	5
3.2.	Realiziran primer uporabe .....	5
4.	FUNKCIONALNE ZAHTEVE.....	7
4.1.	Osnovne zahteve.....	7
4.2.	Vnos, obdelava in integracija podatkov (ETL) .....	7
4.3.	Podatkovno skladišče in podatkovne baze .....	8
4.4.	Zahteve za varnost in nadzor dostopa.....	8
4.5.	Analitika in vizualizacija.....	9
4.6.	Upravljanje, nadzor in optimizacija delovanja platforme .....	9
4.7.	Podatkovni viri .....	10
4.8.	Podatkovna integracija in ETL procesi .....	11
4.9.	Use case – primeri uporabe .....	12
4.10.	Tehnični opis rešitve .....	15

## 2. UVOD

Implementacija in vzpostavitev sistema poročanja, ter izvajanje analiz ELES – projekt »Vzpostavitev BIG Data platforme« mora v enotno okolje integrirati obstoječe vsebine in funkcionalnosti ter jih skladno s potrebami razširiti. Prav tako predvidevamo vpeljavo sistema za urejanje matičnih podatkov in analiziranje ogromnih količin podatkov (Big Data).

S projektom želimo uporabnikom ponuditi sodobno platformo za podatkovno, operativno in poslovno analitiko. Le-ta mora temeljiti na celovitem integriranem enotnem podatkovnem skladišču in analitičnem modelu, do katerega uporabniki lahko dostopajo preko napredne uporabniške rešitve za poročanje in analiziranje. Z uvedbo sistema za poslovno in podatkovno analitiko se vzpostavi centraliziran in poenoten vir za spremljanje poslovanja, poročanje in izvajanje analiz. Takšen vir mora omogočati, da se nad podatki implementira napredne poslovne in operativne scenarije tudi na ogromnih količinah podatkov (Sistem za hrambo in obdelavo masovnih podatkov) tako na zahtevo, kot tudi v realnem času. Postavitev same platforme mora predvidevati tudi vpeljavo sodobnih orodij za vizualizacijo, raziskovanje, samopostrežno analitiko in navsezadnje tudi napredno analitiko.

Projekt je zaradi lažje evaluacije razdeljen na več faz, in sicer implementacija integracijske platforme za zbiranje BigData in produkcijska postavitve celovitega BigData informacijskega sistema (IS), razvoj uporabniških primerov, optimizacija delovanja in nadaljnji razvoj.

Pilotni projekt implementacije integracijske platforme za BigData je omogočil testno delovanje BigData platforme in potrditev za zmanjševanje tveganj pred prehodom v produkcijsko delovanje.

V II. Fazi projekta se preide v produkcijsko delovanje na podlagi pilotnega projekta.

### 3. REZULTATI PILOTNEGA PROJEKTA

V okviru pilotnega projekta je bila vzpostavljena osnovna funkcionalnost »Integracijska platforma za zajem Big Data« s ciljem vzpostavitve dostopov do virov za zagotavljanje zbiranja podatkov iz različnih virov v in verifikacijo na enem primeru uporabe.

#### 3.1. Vzpostavljeni dostopi do virov podatkov

V okviru pilotnega projekta je bila vzpostavljena in instalirana Cloudera BigData platforma z umestitvijo v IS ELES. Za potrebe pilotnega projekta so bile uporabljene časovno omejene (trial) licence za Cloudera CDP platformo.

Vzpostavljeni in preverjeni so osnovni ETL-procesi za zajem in hrambo podatkov iz različnih virov z omejenim naborom podatkov ter zagotovljen dostop do podatkov prek interaktivnega SQL-orodja HUE. Vzpostavitev zajema, hranjenje, obdelavo in vizualizacijo podatkov:

- Vzpostavitev ETL (Extract Transform Load) zajema in hrambe podatkov iz različnih virov z omejenim naborom podatkov:
  - Zajem merilnih podatkov za porabljeno električno energijo, napetosti in tokove iz cca. 1200 pametnih števec z resolucijo 15 min, 1minuta za zgodovino 1 leta Vir: Advance => SQL Server (Batch zajem).
- Zajem SCADA podatkov iz csv datotek za zadnje 3 mesece.
- Zajem SCADA podatkov iz PostgreSQL.
- Zajem podatkov sistemov balancing in trgovalnih pretočnih podatkov (Data streaming) preko protokola MQTT/AMQP z resolucijo 1 minute (ECCO SP in druge tehnologije).
- Zajem podatkov merilnih in zaščitnih sistemov v resoluciji 1 sekunde.

Večstopenjska obdelava podatkov temelji na Pyspark ogrodju, ki omogoča robustne transformacije, agregacije in validacije podatkov ter zapis podatkov v strukture podatkovnih jezer (HDFS/Parquet, Kudu, Iceberg). Razvit je bil osnovni in razširjeni dimenzijski model (BUS matrika, dimenzije in dejstva) za podporo analizi podatkov in razvoju uporabniških primerov. Dimenzijski model omogoča integracijo podatkov iz več virov in pripravo podatkov za generične BI/analitične rešitve (npr. Power BI, Apache Superset).

V okviru pilotskega projekta je bil zasnovan in validiran primer uporabniškega dostopa in vizualizacije z uporabo orodja orodje za ad-hoc poizvedbe prek Apache Hue kot osnova za samopostrežno analitiko, vključno z naprednimi vizualizacijskimi orodji.

#### 3.2. Realiziran primer uporabe

Realiziran primer uporabe na integracijski platformi za zajem Big Data je zajemal aktivacijo, nadzor in optimizacija avtomatske rezervne regulacije frekvence (aRPF).

V tem primeru je bil vzpostavljen end-to-end ETL proces, ki vključuje zajem, transformacije in skladiščenje podatkov iz sistema SCADA (z dvosekundno ločljivostjo) in MOL datotek (trgovalni podatki) ter omogoča modeliranje in analizo podatkov v kontekstu aktivacije aRPF v elektroenergetskem sistemu. V nadaljevanju je podan povzetek izvedenih aktivnosti in rezultatov pilotnega projekta:

Vzpostavljen zajem podatkov:

- neposreden zajem podatkov iz SCADA baze (PostgreSQL) v dvosekundnih intervalih ter iz MOL XML datotek (balansiranje, trgovanja),
- omogočeno upravljanje z aktivacijskimi podatki aRPF, podatki o voznih redih, frekvencah ter sistemskih odstopanjih.

#### Transformacija in shranjevanje:

- surovi podatki se hranijo v strukturiranih tabelah v HDFS (Parquet),
- uporaba Pyspark frameworka za robustno čiščenje, standardizacijo, validacijo in integracijo različnih vrst podatkov (dejstva in dimenzije),
- obdelani podatki omogočajo uporabo za analitiko ali vizualizacijo.

#### Analiza in poročanje:

- realizacija izračuna aktiviranih energij aRPF in stroškov po posameznih območjih in časovnih resolucijah (1-min, 15-min),
- izvedba preverjanja skladnosti MOL podatkov in analiza izravnave sistema ob različnih aktivnostih (hitre aktivacije, ročni posegi),
- podatki in rezultati so dostopni za analitiko in napredne BI rešitve (Power BI, Superset).

#### Validacija in zagotavljanje podatkovne kakovosti:

- vzpostavljena je bila validacija in dopolnjevanje manjkajočih podatkov ter preverjanje skladnosti in transparentnosti vseh integracij,
- realiziran dimenzijski model je bil omejen na 10 dimenzij (5 kompleksnih, 5 enostavnih) in do 7 metrik, ter omogoča nadaljnje razširitve z novimi viri, metapodatkovnimi strukturami in povečanje kompleksnosti analitik.

Rezultati pilotskega projekta omogočajo nadgradnjo na več primerov uporabe in razširjene zajeme podatkov.

## 4. FUNKCIONALNE ZAHTEVE

V nadaljevanju predstavljamo funkcionalne zahteve implementirane rešitve.

### 4.1. Osnovne zahteve

R 1. Big Data platforma mora omogočati:

- Visoko skalabilno infrastrukturo.
- Zajem in obdelavo podatkov v realnem času.
- Zanesljivo podatkovno hrambo z več nivoji.
- Napredne analitične in AI/ML zmožnosti.
- Varnost, skladnost in nadzor, skladno z zahtevami energetskega sektorja.
- Samopostrežni dostop in podpora za podatkovne ekipe.
- Data Engineering - Zajem, integracija, priprava in orkestracija podatkovnih tokov (ETL/ELT).
- Streaming / Data Flow - Real-time obdelava podatkov, streaming analytics, podatki iz senzorjev, dogodkov.
- Data Warehouse & Lakehouse - Analitika na velikih količinah podatkov, hrambe podatkov in poizvedbe, podpora za jezero podatkov (data lake) + skladišče (warehouse).
- Operational Database - Sistemi za operativne aplikacije s skalabilno podatkovno bazo v realnem času.
- Machine Learning / AI - Okviri in okolje za data science, modeliranje, strojno učenje, uvajanje modelov v produkcijo.
- Unified Governance & Security (SDX) - Enoten metapodatkovni katalog, upravljanje varnosti, politika dostopa, lineage podatkov.
- Fleksibilnost infrastrukture: možnost uporabljati on-premise, javni oblak ali hibridno arhitekturo. Primerno, če imate kombinacijo sistemov in želite minimalizirati migracijske zaprte poti.
- Skalabilnost in visoke količine podatkov: Platforma je zasnovana za petabajte-razred, torej lahko ustreza zahtevam "big data".
- Odprti standardi in manj vendor-lock-ina: Ker temelji na odprtokodnih tehnologijah in omogoča uporabo različnih skladišč/virov.
- Celosten pokrivalni pristop: Od zajema podatkov, preko obdelave, skladiščenja, analitike do ML in AI – dobra osnova za "end-to-end" big data platformo.
- Varovanje, upravljanje in skladnost: Z integriranim metapodatkovnim slojem in politiko za varnost, kar je ključno za regulirane sektorje (npr. energetika).
- Podpora za real-time in več-okolja (multi-cloud/hybrid): Kar je pogosto ena ključnih zahtev v modernih arhitekturah.

### 4.2. Vnos, obdelava in integracija podatkov (ETL)

Arhitektura BigData platforme mora zagotavljati tako podporo za neprekinjen zajem podatkov v realnem času kot tudi paketno obdelavo večjih količin podatkov iz različnih virov. Z integracija naprednih orodij za podatkovni inženiring mora biti omogočeno zanesljivo in skalabilno izvajanje ETL procesov s podporo za različne podatkovne tokove, transformacije in optimizirano pripravo podatkov za nadaljnjo analizo in uporabo v poslovnih procesih. Minimalne zahteve za vnos, obdelavo in integracijo podatkov (ETL) so naslednje:

R 2. Podpora za realnočasovni (streaming) in paketni (batch) zajem podatkov iz različnih virov (npr. SCADA, PSI, SUMO, NMM, VR, MX, D365, itn.)

- R 3. Standardizirani ETL vmesniki za zajemanje, replikacijo in transformacijo podatkov iz omenjenih virov v podatkovno platformo.
- R 4. Platforma mora zagotavljati orodja za podatkovni inženiring za izvedbo ETL procese na osnovi pySpark in Java.

### 4.3. Podatkovno skladišče in podatkovne baze

Podatkovno skladišče za BigData platformo mora zagotavljati robustno, razširljivo in visoko zmogljivo okolje za shranjevanje, upravljanje in analiziranje podatkov. Sistem mora podpirati sodobne rešitve za Big Data podatkovna skladišča na osnovi platform kot so Iceberg, HIVE, Kudu in HDFS, kar omogoča učinkovito obdelavo kompleksnih podatkovnih nizov in podporo analitičnim potrebam družbe ELES d.o.o. Minimalne zahteve:

- R 5. Big Data skladišče: podpora za DW, CDW, Iceberg, HIVE, Kudu, HDFS..

### 4.4. Zahteve za varnost in nadzor dostopa

Za zagotavljanje celovite varnosti in upravljanja dostopa do podatkov, mora BigData platforma vključevati napredne mehanizme za avtentikacijo, avtorizacijo ter upravljanje uporabniških in administrativnih pravic. Implementacija mora omogočati centralizirano upravljanje varnostnih pravil ter nadzor dostopov na ravni datotek, tabel in stolpcev, vključno s sledenjem in revizijo vseh dostopov za zagotavljanje skladnosti z regulativnimi zahtevami. Integracija z zunanjimi sistemi za upravljanje identitet (Microsoft Active Directory oziroma Entra ID) mora zagotavljati enotno in zanesljivo upravljanje uporabnikov ter krepiti robustnost avtentikacije. Ključno je, da platforma zagotavlja zaščito podatkov po najnovejših varnostnih standardih, vključno s šifriranjem ter zaščito pred nepooblaščenimi dostopi in omogočati učinkovito upravljanje podatkovnih katalogov, governance ter metapodatkov za transparentnost in sledljivost podatkovnega ekosistema. Minimalne zahteve za varnost in nadzor dostopa:

- R 6. centralizirano upravljanje pristojnosti, avtentikacija in avtorizacija vseh uporabnikov ter dostopov do podatkov in storitev,
- R 7. centralizirano določanje in izvajanje varnostnih pravil na nivoju datotek, tabel in stolpcev z zmožnostjo revizije vseh dostopov za potrebe skladnosti z regulative,
- R 8. integracija s sistemom za upravljanje identitet (npr. MS Active Directory oz. Entra ID) za enotno upravljanje uporabniških računov in pravic,
- R 9. implementacija naprednih avtentikacijskih mehanizmov, ki preprečujejo nepooblaščen dostop do platforme, podatkov in storitev,
- R 10. zagotavljanje zaščite podatkov skozi celoten življenjski cikel s pomočjo šifriranja, nadzorovanih dostopov in spremljanja vseh aktivnosti na podatkih,
- R 11. upravljanje podatkovnih katalogov, vodenje metapodatkov, data governance in zagotavljanje popolne sledljivosti ter transparentnosti nad podatkovnimi viri in tokovi..



## 4.5. Analitika in vizualizacija

Za podporo podatkovno podprtemu odločanju, napredni analitiki in učinkoviti poslovni uporabi mora Big Data platforma zagotavljati celovito okolje za izvajanje naprednih analitičnih procesov, strojnega učenja ter interaktivno vizualizacijo podatkov. Ključna je tudi integracija z različnimi analitičnimi in vizualizacijskimi orodji ter intuitivni uporabniški vmesniki, ki končnim uporabnikom omogočajo neposreden dostop do podatkov, izvajanje poizvedb in gradnjo vizualnih vpogledov glede na lastne potrebe. Minimalne zahteve za analizo in vizualizacijo so naslednje:

- R 12. integracija z vodilnimi orodji za vizualizacijo in poročanje (npr. Apache Superset, MS Power BI, Grafana, IBM Cognos), ki omogočajo enostavno ustvarjanje nadzornih plošč, poročil in različnih podatkovnih vizualizacij,
- R 13. omogočen dostop do podatkov z uporabo SQL (Impala) ter interaktivnimi uporabniškimi vmesniki za izvajanje poizvedb, upravljanje podatkovnih tokov in osnovne vizualizacije,
- R 14. intuitiven, interaktiven uporabniški vmesnik, ki podpira učinkovito izvajanje in vizualizacijo kompleksnih analiz ter omogoča prilagodljivo upravljanje pravic dostopa na ravni posameznih uporabnikov ali skupin za granularno kontrolo vizualizacijskih funkcionalnosti,
- R 15. zagotovljeni so API-ji za razvoj prilagojenih aplikacij in rešitev za specifične potrebe posameznih uporabnikov ali poslovnih procesov,
- R 16. podpora za napredne analitične in AI procese, vključno z obdelavo podatkov, razvojem in uporabo algoritmov strojnega učenja (Machine Learning), orodji, kot sta Jupyter in Python.

## 4.6. Upravljanje, nadzor in optimizacija delovanja platforme

Za zagotavljanje visoke razpoložljivosti, stabilnosti ter učinkovitosti delovanja Big Data platforme je potrebno celovito upravljanje tako nadzornih kot strojnih virov. Platforma mora omogočati neprekinjen nadzor in upravljanje nad vsemi ključnimi komponentami, avtomatizirane mehanizme za replikacijo, varnostno kopiranje in nadomestne načine delovanja ter dinamično prilagajanje razpoložljivih strojnih virov glede na aktualne obremenitve in potrebe uporabnikov. Minimalne zahteve za upravljanje, nadzor in optimizacija delovanja platforme so naslednje:

- R 17. zagotovitev avtomatiziranih mehanizmov za replikacijo podatkov, varnostno kopiranje (backup) in hitro obnavljanje sistema v primeru izpadov ali napak,
- R 18. centraliziran nadzor in upravljanje vseh storitev, aplikacij, uporabnikov in dostopa prek enotnega kontrolnega vmesnika (Replication Manager, Supervision, Management Console),

- R 19. podpora za dinamično dodeljevanje procesorskih in pomnilniških virov različnim aplikacijam z namenom doseganja optimalne izkoriščenosti in stabilnosti sistema,
- R 20. omogočena visoka stopnja prilagodljivosti na spremembe v delovnih obremenitvah, vključno s samodejno porazdelitvijo nalog po vozliščih za povečanje zmogljivosti platforme,
- R 21. integracija rešitev za centralizirano upravljanje strojnih virov, ki omogoča stalni monitoring, optimizacijo porabe virov ter enostavno povezovanje z aplikacijami, ki izvajajo porazdeljeno procesiranje podatkov,
- R 22. zagotovljena mora biti zanesljivost, varnost in visoka razpoložljivost vseh komponent platforme tudi v primeru nepredvidenih obremenitev ali tehničnih težav.

## 4.7. Podatkovni viri

Ponudnik mora zagotoviti celovito integracijo in zajem raznovrstnih podatkovnih virov v Big Data, vključno z merilnimi podatki, geografskimi in tehničnimi podatki, poslovnimi in operativnimi sistemi ter eksternimi podatki. Vsi podatkovni viri morajo biti strukturirani, ažurni in dostopni v skladu s tehničnimi zahtevami naročnika ter omogočati učinkovito analizo, optimizacijo elektroenergetskega omrežja, izboljšanje poslovnih procesov in podporo pri strateškem odločanju.

V nadaljevanju so opisani posamezni podatkovni viri, njihova vloga v Big Data okolju ter ključne tehnične zahteve za njihovo integracijo in obdelavo.

V nadaljevanju so podani podatkovni viri, ki se v okviru projekta integrirajo na BigData platformo (popis podatkovnih virov ELES):

- R 23. SCADA
- R 24. PQ (MKEE) – Meritve kakovosti napetosti
- R 25. Protel – Sistem za pregled zajemov meritev in delovanj zaščitnih naprav
- R 26. SDM – Storitve je namenjena avtomatskemu zajemu dogodkov zaščitnih naprav
- R 27. ADVANCE – Števnice meritve
- R 28. GIS – Geografski IS
- R 29. Maximo – upravljanje s sredstvi
- R 30. SCALAR – Atmosferske razelektritve
- R 31. OTLM – SUMO na transformatorjih,..
- R 32. DAC SCADA – črpanje podatkov iz OT okolja
- R 33. DAC IoT – zajem podatkov IoT platform

## 4.8. Podatkovna integracija in ETL procesi

Big Data platforma mora vsebovati komponente za podatkovno integracijo in obdelavo, ki morajo biti zasnovane na način, ki omogoča zajem, transformacijo in dostop do podatkov iz različnih virov, pri čemer mora zagotavljati visoko stopnjo avtomatizacije, prilagodljivosti in skladnosti s potrebami naročnika. Cilj integracije in obdelave podatkov je zagotoviti enotno podatkovno okolje za analitične in operativne namene, ki bo omogočalo učinkovito podporo pri sprejemanju poslovnih odločitev.

Podatkovna integracija je ključna funkcionalnost Big Data platforme, ki mora omogočati zajem podatkov iz raznolikih virov, vključno s strukturiranimi podatki, kot so relacijske baze in datoteke v formatih CSV, XML ali JSON, ter nestrukturiranimi podatki, kot so besedilne datoteke, slike, logi in dogodki. Poleg tega mora platforma podpirati pretoke podatkov v realnem času z uporabo protokolov, kot so MQTT, REST API in MQ, ter omogočati zajem podatkov iz naprav IoT in naprednih merilnih sistemov preko standardnih protokolov, kot sta MQTT in HTTP. Za učinkovito obdelavo podatkov mora platforma podpirati zajem podatkov v realnem času z minimalno zakasnitvijo, kar omogoča proaktivno ukrepanje, ter obenem uporabo napredne analitike za masovno obdelavo velikih količin zgodovinskih podatkov. Poleg tega mora omogočati prenos podatkov preko različnih protokolov, vključno s FTP/SFTP, SOAP in REST API, s čimer zagotavlja fleksibilnost pri integraciji z različnimi podatkovnimi viri in aplikacijami.

Platforma mora vključevati zmogljive ETL procese za zajem, procesiranje in transformacijo podatkov v dimenzijske modele.

- R 34. Zajem podatkov mora biti omogočen iz različnih virov v njihovi surovi obliki ter shranjevanje v podatkovno jezero (ang. Data Lake) z uporabo naprednih orodij za integracijo, ki podpirajo grafične vmesnike za enostavno konfiguracijo procesov. Ob enem mora biti fleksibilen in omogočati proženje procesov na podlagi urnika, dogodkov ali sprememb v podatkovnih virih (ang. event-driven processing).
- R 35. ETL procesi morajo zagotavljati zajem, standardizacijo, čiščenje in pripravo podatkov za nadaljnjo uporabo. Za masovno procesiranje in hitro transformacijo podatkov mora platforma uporabljati komponento Apache Spark 3.2 ali novejša verzija, ki omogoča visokozmogljivo obdelavo podatkov na več vozliščih v grozdu in je primerna za velike količine podatkov.
- R 36. Procesji podatkovnih obdelav, morajo biti zasnovani tako, da omogočajo razširitev in prilagoditev glede na prihodnje potrebe, kot so vključitev novih podatkovnih virov ali spremembe v obstoječih procesih. Prilagodljivost mora podpirati hitro uvajanje novih funkcionalnosti brez motenj v delovanju sistema.
- R 37. ELT procesi morajo omogočati paralelno obdelavo podatkov, kar vključuje obdelavo velikih količin zgodovinskih podatkov in gradnjo kompleksnih dimenzijskih modelov, ki vključujejo tako aktualne kot zgodovinske podatke.

R 38. Grafični vmesnik za konfiguracijo procesov mora omogočati enostavno definiranje toka podatkov z vizualno predstavitevjo posameznih korakov procesiranja, vizualizacijo posameznih stopenj obdelave, kar omogoča lažje sledenje procesom in hitro diagnosticiranje težav ter proženje procesov na podlagi vnaprej določenih pravil, kar zagotavlja avtomatizacijo in optimizacijo delovnih tokov.

## 4.9. Use case – primeri uporabe

Popis Uporabniških primerov, ki jih je potrebno vpeljati v Big data platformo:

R 39. UC 1 – Spremljanje delovanja zaščitnih relejev

- Ob dogodku na sistemu se shrani COMTRADE datoteka z natančnim popisom dogodka. Datoteke so shranjene v tehničnem omrežju. Podatkovni vir je SDM600, ki zbira podatke iz vseh relejev, iz vsake strani elementa po eno datoteko in ker je podvojen sistem je vse x2
  - i. COMTRADE datoteke imajo podatek tudi o elementu, ki ga je generiral (recimo pot, kjer je datoteka, serijska številka,..).
  - ii. Zahtevana je uskladitev s SCADA sistemom, da se bo lahko ločilo testne od dejansko pomembnih COMTRADE datotek / dogodkov.
  - iii. Te datoteke se nahajajo v OT segmentu.
  - iv. COMTRADE datoteke so zajeti oscilogrami toka, ko pride do izpada in se zapisujejo samo, ko je zaščitni rele v delu.
  - v. Vsebino oscilograma se mora ob zajemu oziroma zapisu v BigData okolje razčleniti v tabelarično obliko.
- Ostali viri podatkov, ki so potrebni za izvedbo analiz:
  - vi. Maximo: Podatki o sredstvih - maksimalen tok.
  - vii. SCADA podatki so vir podatkov o stanju ob dogodku.
  - viii. Tabelarični podatki EXCEL so vir podatkov o pričakovanih kratkostičnih tokovih v vseh stikališčih in jih je potrebno oblikovati kot nov centralni vir podatkov (ni vir v BigData). Pričakovani kratkostični tokovi predstavljajo najvišje kratkostične tokove po postajah (odklopnik v stikalnem polju)
- Tako SDM600 kot Protel imata meta podatke o elementu, ki je generiral COMTRADE datoteko.
- Vizualizacija mora obsegati prikaz kratkostičnih tokov ob dogodku, prikazan mora biti tudi potek kratkostičnega toka znotraj omejitev in primerjava s pričakovanimi kratkostičnimi tokovi. Vizualizacija mora biti izdelana v Power BI, pripravi se možnost izdelave vizualizacij v Superset.

R 40. UC 2 – Zajem podatkov StatCom Fault Recorder (Static Synchronous Compensator)

- StatCom Fault Recorder, zapisuje:
  - ix. COMTRADE datoteke ob napakah,
  - x. 10min povprečja napetosti.

- StatCom je priključen v eni točki omrežja na 400 kV, težave so s prevelikimi napetostmi.
- Uporabniški primer predstavlja zajem podatkov v BigData okolje, ki se sedaj ročno pobirajo z USB ključki iz same naprave.
- Podatki se morajo zapisati v BigData okolje najmanj 1x dnevno.

R 41. UC 3 – Zbiranje Power Quality (PQ) podatkov – POC

- To so tri različne naprave - Dranetz, Iskratel, Eberle (tega je največ).
- Obstajajo 10-min vrednosti, na on-error dogodek se naredi oscilografija.
- PQ podatki vsebujejo: odstopanje napetosti, frekvenčna stabilnost, THD, faktor moči.
- Podatki morajo biti iz vseh treh tipov/sistemov združeni v Big Data platformi.

R 42. UC 4 – POS/SRT – Aktivacija aRPF, nadzor in optimizacija

- V BigData platformo se mora uvoziti:
  - i. iz sistema SCADA podatke v 2sec resoluciji,
  - ii. podatke o aktiviranih količinah aRPF,
  - iii. vse ostale podatke za izračun aktiviranih količin aRPF (vozni red, frekvenca, odstopanje sistema...),
  - iv. MOL datoteke, ki jo uporabi SCADA za aktivacijo aRPF,
  - v. MOL podatke iz Balancing Clienta.
- Na podlagi tega se v BigData platformi zagotovi sledeče:
  - i. Izračun aktiviranih energij aRPF po posameznih PSI in stroški aktivacij v minutni in 15minutni resoluciji,
  - ii. preveri se ali so bili v SCADA sistemu ob aktivacijah uporabljeni enaki MOL podatki, kot jih imamo istočasno v Balancing sistemu,
  - iii. ob aktivacijah hitrih Pent in vnosu ročnih odmikov operaterjev v SCADA se mora preveriti pravilnost izravnave sistema (prikaz oz. analiza v 2sec resoluciji)

R 43. UC 5 – DAC–Zajem vsebin SSS

- dostop do vsebin z uporabo webservice-a,
- zajem podatkov bo v neki razumni frekvenci, običajno dnevno, po potrebi urni zajem,
- Inicijalni zajem podatkov se opravi z kopijo baze oziroma se zajem izvede po mesecih.

Predlog rešitve sta dva ETL-ja, en za zajem, drugi za popravljanje podatkov.

- Cca. 1200 merilnih mest, števci so proizvajalcev Iskra Emeco, CV in Landis.
- Nekatera merilna mesta imajo minutno periodo, v okviru projekta se zajema samo merilna mesta s 15-minutno periodo
- Zajem zgodovine se bo izvedel za mesečne količine po fazah skozi webservice, da ne obremenimo produkcije

- Količina podatkov 1,2TB je celotna baza za 10+let (vse) (CLSQLmeritve\SLQMERITVE, letni prirast je 60GB na leto, letno se loči tablespace
- Obstaja tudi testni sistem brez Web servisa
- Tadej Pripravi specifikacijo/strukturo web service-a
- ID merilnega mesta je povezan z Assetspac\AssetNo OZN\_MM v MX, posledično bo potreben zajem podatkov za Asset in AssetSpec iz MX
- Potrebno je upoštevati spremembo naziva merilnega mesta, ker je verjetno mogoče identificirati v MX
- Podatki, ki se prenašajo near-real-time, so lahko tudi manjkajoči, kar je potrebno upoštevati in korigirati v naslednjih prenosih, recimo v tridnevni periodah. Podatki se vedno popolnijo.
- Tadej preveri še čas, ki je zapisan v bazi in kak je v WebService
- WebService dobi le delto/spremembo moči iz prejšnje 15min meritve
- Števci se na 6 let preverjajo/umerjajo, serijska številka števca ni v podatkih, ampak je le v MX. 20 % števec se letno umeri.
- V kolikor bo potreba, se lahko pridobijo tudi minutni podatki iz parih števec. Zaenkrat se bodo prenašali samo 15 minutni.
- Kvaliteta prenesenih podatkov je ok, čiščenje ni potrebno
- Pridobivali se bodo samo števrni podatki (delta), A+, A-, R+ in R-, ter timestamp in deviceid Podatki o napravah (devicedetails) in podobni se ne bodo prenašali. Govoreče kode. Ni dnevnih stanj.

#### R 44. WAMS

- zajem podatkov iz PMU merilnikov, cca 150 naprav z resolucijo 50 vzorcev/sec, merijo, fazorje napetosti in toka po fazah ter frekvenco in spremembo frekvence (ROCOF), izračunajo je P, Q, simetrične komponente, in podobno. Ključni parametri uporabniškega primera:
  - Uporablja se ELPROS elektronski in programski sistemi d.o.o. - SW WAProtector™,
  - WAProtector™ uporablja interno binarno krožno 30 dnevno bazo (podatki so samo na zahtevo, CSV, Comtrade, ),
  - Količina podatkov je cca 6TB/30dni, 7.5GB/30dni na en feeder,
  - Uporablja se IEEE C37.118 protokol - stream podatkov med PMU merilniki in PDC-jem, za nadaljnje deljenje podatkov z ostalimi deležniki se uporablja protokol IEC104,
  - neto prirast podatkov za pol leta je 6T.

## 4.10. Tehnični opis rešitve

Ponudnik mora ob oddaji ponudbe za pridobitev točk iz naslova tega merila (Ad 2) ponudbi predložiti tehnični opis rešitve s sledečimi vsebinami:

- R 45. Organizacijo projekta.
- R 46. Koncept vzpostavitve rešitve.
- R 47. Identifikacija ključnih komponent in časovnih tveganj projekta.
- R 48. Koncept izvedbe projekta.
- R 49. Koncept za nadaljnji razvoj zajema podatkov in implementacijo napredne analitike.