

# Technical Documentation

Misi 2 kelompok 16

## 1. Executive Summary (Pendahuluan)

Dokumen ini menjelaskan desain teknis dan implementasi Data Warehouse untuk Satuan Pengawas Internal (SPI). Sistem ini dirancang untuk mengintegrasikan data dari Sistem Internal Audit, HRIS, dan Sistem Organisasi guna menyediakan "Single Source of Truth" bagi analisis risiko, kinerja auditor, dan status tindak lanjut temuan audit.

### 1.1 Ruang Lingkup Data

- **Fact Grain:** Satu baris mewakili satu **Temuan Audit** per **Rekomendasi**.
- **Partisi Data:** Data dipartisi berdasarkan Tahun Audit (Range 2020-2025) untuk optimalisasi performa .

## 2. System Architecture

Sistem menggunakan pendekatan **ELT (Extract, Load, Transform)** dengan arsitektur database sebagai berikut:

1. **Staging Area (stg schema):** Tempat penampungan data mentah dari sumber sebelum transformasi.
2. **Data Warehouse (dbo schema):** Model dimensional teroptimasi (Star Schema).
3. **Orkestrasi ETL:** Menggunakan T-SQL Stored Procedures (`usp_Load_Dim_*` dan `usp_Load_Fact_*`) yang dibungkus dalam transaksi.

## 3. Dimensional Modeling (Logical Design)

### 3.1 Skema Model

Menggunakan **Star Schema** dengan satu tabel fakta pusat yang dikelilingi oleh tabel dimensi.

### 3.2 Kamus Data (Data Dictionary)

#### A. Tabel Fakta: Fact\_Temuan\_Rekomendasi

Menyimpan metrik kuantitatif audit.

Nama Kolom	Tipe Data	Keterangan	Aturan Bisnis
Fakta_SK	BIGINT	Primary Key	Surrogate Key (Identity).
Waktu_SK	INT	Foreign Key	Format integer YYYYMMDD.
Tahun_Audit	INT	<b>Partition Key</b>	Degenerate Dimension untuk partisi.
Auditor_SK	INT	Foreign Key	Relasi ke Dim_Auditor.
Unit_Kerja_SK	INT	Foreign Key	Relasi ke Dim_Unit_Kerja.
Jumlah_Temuan	INT	Measure	Default 1.
Skor_Risiko	DECIMAL	Measure	Rentang 1.00 - 5.00.
Potensi_Kerugian	BIGINT	Measure	Nilai Rupiah temuan.
Usia_Rekomendasi	INT	Derived Measure	Aging (Hari keterlambatan).

## B. Tabel Dimensi Utama (SCD Type 2)

Fokus pada dimensi yang memiliki perubahan historis (Slowly Changing Dimension).

## 1. Dim\_Auditor

- **Natural Key:** ID\_Sistem\_Sumber (NIP/ID Asli).
- **SCD Type 2 Attributes:** Jabatan, Tim\_Audit.
- **Tracking Columns:** EffectiveDate, ExpiryDate, IsCurrent (1=Aktif, 0=Histori).

## 2. Dim\_Unit\_Kerja

- **Natural Key:** Kode\_Unit.
- **SCD Type 2 Attributes:** Jenis\_Unit, Kepala\_Unit.

## 4.1 Partitioning Strategy

Untuk menangani volume data besar, tabel fakta dipartisi secara horizontal berdasarkan tahun.

- **Partition Function:** PF\_AuditYear (Range Right).
- **Boundaries:** 2020, 2021, 2022, 2023, 2024, 2025 .
- **Partition Scheme:** PS\_AuditYear memetakan seluruh partisi ke Filegroup PRIMARY .

## 4.2 Indexing Strategy

Strategi indexing multi-layer diterapkan untuk berbagai kebutuhan query:

1. **Clustered Columnstore Index:** Diterapkan pada data analitik (Hot Data) dengan filter WHERE Tahun\_Audit > 2023 untuk agregasi super cepat .
2. **Clustered Index (Partitioned):** CIX\_Fact\_Temuan\_Waktu pada (Waktu\_SK, Fakta\_SK) selaras dengan skema partisi.
3. **Non-Clustered Indexes:** Covering index pada kolom Foreign Key (Unit, Auditor) dengan klausa INCLUDE untuk menghindari *Key Lookup*.

## 5. ETL Design & Mapping

### 5.1 Source-to-Target Mapping

Berikut adalah logika transformasi utama yang diterapkan dalam Stored Procedure:

Tabel Sumber	Kolom Sumber	Tabel Target	Kolom Target	Logika (T-SQL)	Transformasi

stg.Auditor	Nama_Lengkap	Dim_Auditor	Nama_Auditor	UPPER(TRIM(s>Nama_Lengkap)).
stg.Temuan	Tanggal	Fact_Temuan	Waktu_SK	CAST(CONVERT(VARCHAR(8), s.Tanggal, 112) AS INT).
stg.Temuan	Tgl_Target	Fact_Temuan	Usia_Rekom	DATEDIFF(DAY, s.Tgl_Target, GETDATE()).
stg.Temuan	Tgl_Temuan	Fact_Temuan	Tahun_Audit	YEAR(s.Tanggal_Temuan) (Untuk kolom partisi).

## 5.2 SCD Type 2 Logic (Stored Procedure)

Logika penanganan perubahan dimensi diimplementasikan dalam usp\_Load\_Dim\_Auditor:

- Expire:** Update IsCurrent = 0 dan ExpiryDate = GETDATE() untuk data lama jika Jabatan atau Tim\_Audit berubah .
- Insert:** Insert data baru dengan IsCurrent = 1 dan EffectiveDate = GETDATE() .

## 6. Data Quality Assurance (DQA)

Kualitas data divalidasi secara otomatis menggunakan usp\_Run\_DQ\_Checks yang mencatat hasil ke tabel DQ\_Audit\_Log.

**Rules yang Diterapkan:**

- Completeness:** Memastikan Auditor aktif memiliki NIP, Nama, dan Jabatan (Tidak boleh NULL)
- Consistency (Referential Integrity):** Cek "Orphan Records" di tabel fakta (Fakta tanpa Auditor yang valid di dimensi) .
- Accuracy:** Validasi Skor\_Risiko\_Temuan harus dalam rentang 1.00 - 5.00 .
- Uniqueness:** Mencegah duplikasi data fakta berdasarkan kombinasi Temuan\_SK dan Rekomendasi\_SK .

## 7. Performance Benchmarking

Pengujian dilakukan untuk memvalidasi efektivitas strategi indexing dan partitioning.

Skenario Pengujian	Deskripsi	Target	Hasil Aktual	Index Terpakai
<b>1. Risk Profile Aggregation</b>	Join Unit Kerja & Waktu + SUM Potensi Kerugian <sup>36</sup> .	< 1 dtk	<b>0.05 dtk</b>	NCCIX_Fact_Analytic (Columnstore).
<b>2. Monthly Trend Analysis</b>	Grouping per Tahun/Bulan (Partition Elimination) <sup>38</sup> .	< 2 dtk	<b>0.08 dtk</b>	CIX_Fact_Temuan_Waktu.
<b>3. Auditor Drill-down</b>	Filter spesifik kinerja auditor aktif (IsCurrent=1) <sup>40</sup> .	< 3 dtk	<b>0.12 dtk</b>	IX_Fact_Auditor_Key.

**Kesimpulan:** Sistem telah memenuhi seluruh persyaratan fungsional dan non-fungsional. Implementasi **Partitioning per Tahun** dan **Columnstore Index** terbukti meningkatkan performa query analitik secara signifikan, dengan waktu eksekusi rata-rata di bawah 0.2 detik untuk jutaan baris data.