

Proposed Solution – Active + Archive Tier

We will implement a **two-tier architecture**:

1. **Active Tier** (Cosmos DB):

- Holds only recent 3 months of billing records.
- High throughput (RU/s) optimized for frequent reads and writes.

2. **Archive Tier** (Azure Blob Storage + Azure Function/Logic App):

- Stores historical data older than 3 months as **compressed JSON files** or **parquet format**.
- Accessed via a **transparent data access layer** (function/middleware) that fetches from Blob if the record isn't in Cosmos.

This approach minimizes Cosmos storage costs while keeping old data accessible with seconds-level latency.

Architecture Diagram

