Data Architecture Assessment

(LTL (Less-Than-Truckload) shipping services Logistics company)

Assessed current data architecture of client that has multiple data sources in different platforms and recommended alternate architecture options (On-Premises and Cloud) to improve efficiency, scalability and automation

Data Architecture Assessment For A Logistics Company

ABOUT THE CLIENT

Client provides **LTL (Less-Than-Truckload) shipping services to businesses** across all major metro cities in the U.S.

SITUATION

- Client has been facing **difficulties in maintaining and leveraging the data from multiple systems** for their operations due to older data systems and lack of integration across different data sources
- Merilytics partnered with the client to **assess their current data architecture and recommend a robust (on-premises and cloud-based) data flow** to improve efficiency, scalability and automation

VALUE ADDITION

- Conducted **discussions with data owners** of all data systems and software to understand the current data flow. Developed a **comprehensive questionnaire** for these discussions to capture details such as server information, data flow, replication process, data size, frequency of ingestion etc.
- **Documented the data flow** of all systems, servers and applications; and created **data flow diagram for the current architecture**
- **Assessed the current architecture** to identify the opportunities and gaps for enhancements
- **Recommended an alternate on-premises architecture** to **improve efficiency, scalability, automation** of processes in existing architecture and integration of Ops systems
- **Recommended a cloud-based data architecture** along with a plan to transition from on-premises to cloud environment

IMPACT

- Based on the recommendations, client was able to **bridge the gaps in current architecture** and **integrate all data sources** into a consolidated server, which further **helped in creating an efficient automated BI reporting suite**
- The proposed architecture also **helped client to establish transparent data flows and strong data governance**
- Client also **planned a long-term transition from on-premises to cloud-based data environment** based on the recommendations

Approach & Methodology

Collaboration with Data Owners

Conducted discussions with data owners of all systems and software to understand current data flow

Assessment of Current Architecture

Assessed current data flows and architecture design to identify gaps and opportunities for improvement

Proposed Cloud based Architecture

Designed and proposed an alternate cloud-based architecture for better scalability along with a migration plan from on-premises to cloud-based environment

Documentation of Current Architecture

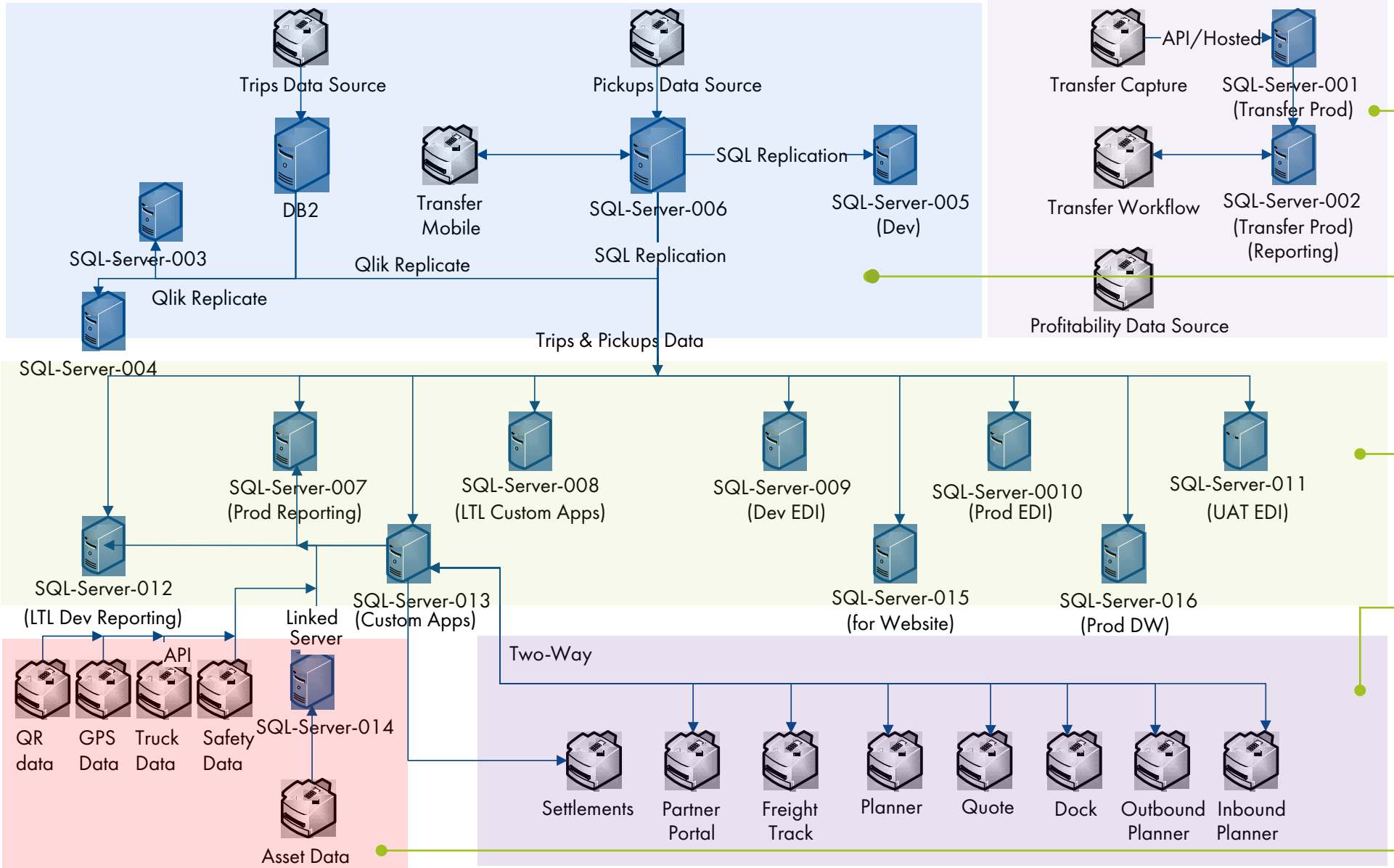
Documented all observations from the initial discussions with data owners and created consolidated data flow diagrams across all systems

Proposed enhanced On-Premises Architecture

Designed and proposed an enhanced on-premises architecture with better efficiency and integration

Existing Data Flow

ILLUSTRATIVE



Some data sources are not integrated to the SQL server environment

Key Logistics data sources are **hosted on older data systems** and the data is replicated to On-premises SQL servers

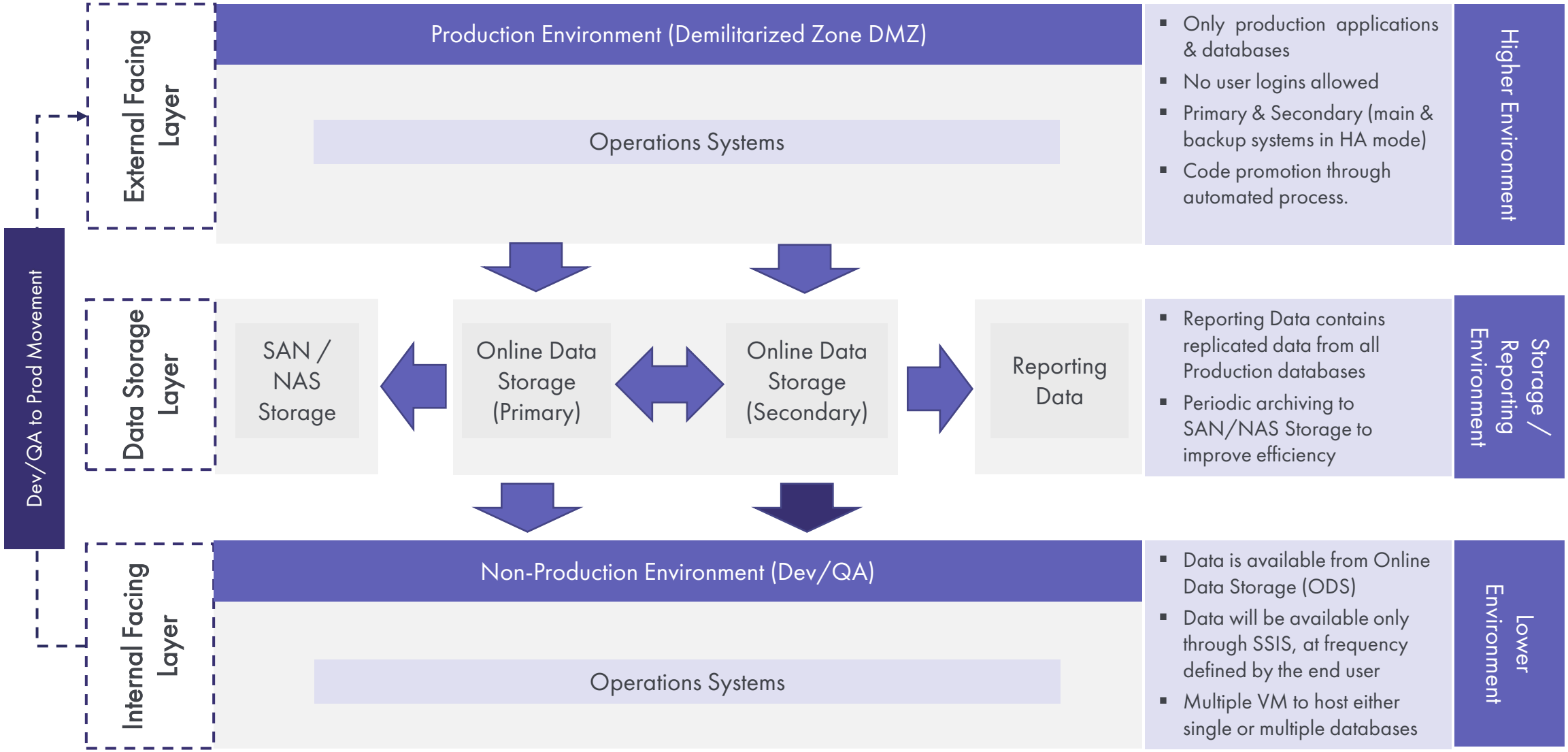
SQL Server Layer: Data from sources is again replicated to multiple SQL servers for different purposes such as reporting, custom applications, website etc.

Custom Apps: A single SQL server acts as the source for multiple internal custom applications built by client for their daily operations

External data sources: Data from external sources is loaded to the reporting servers through multiple API calls

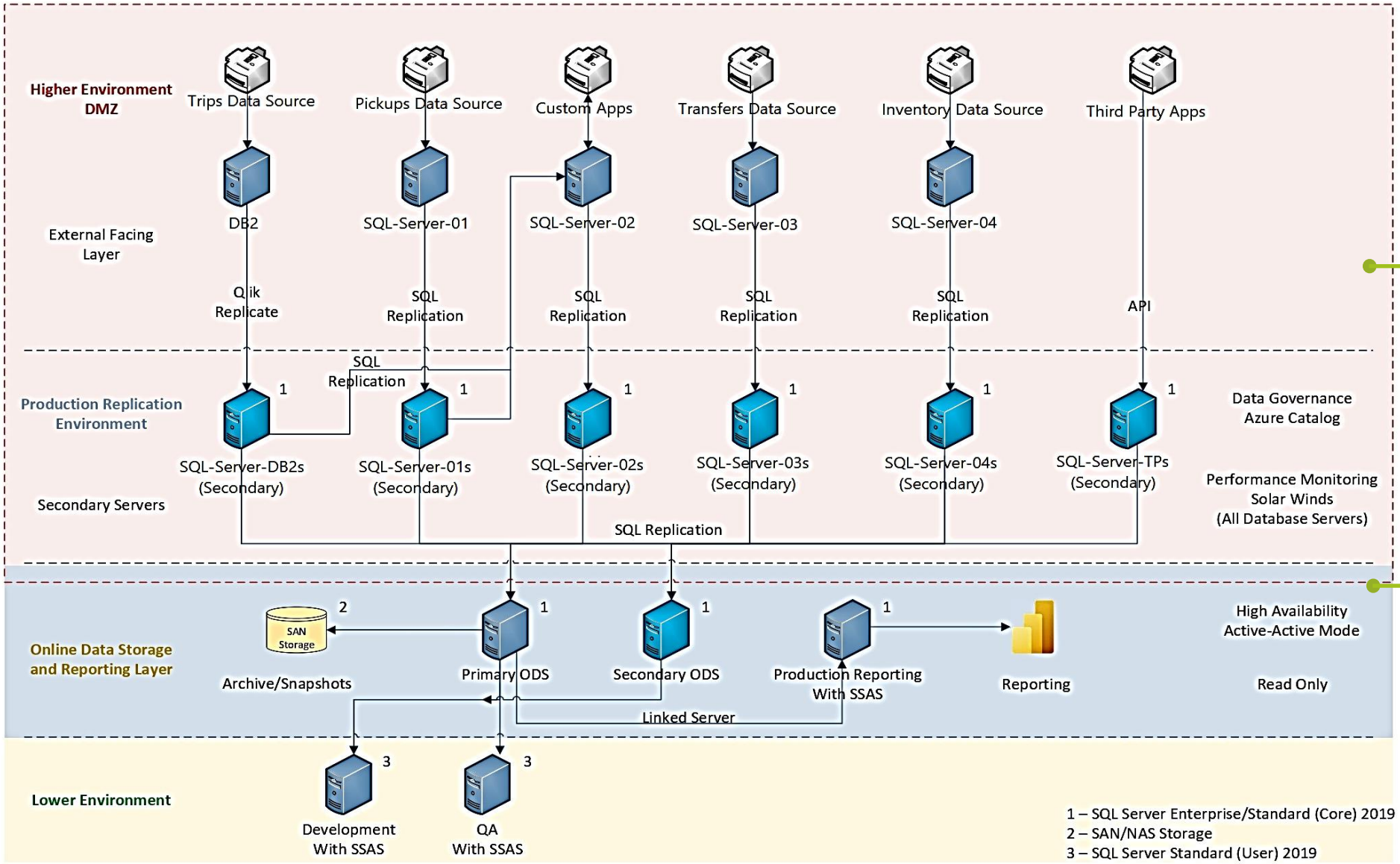
Proposed Data Architecture Framework

ILLUSTRATIVE



Recommended On-premises Network

ILLUSTRATIVE



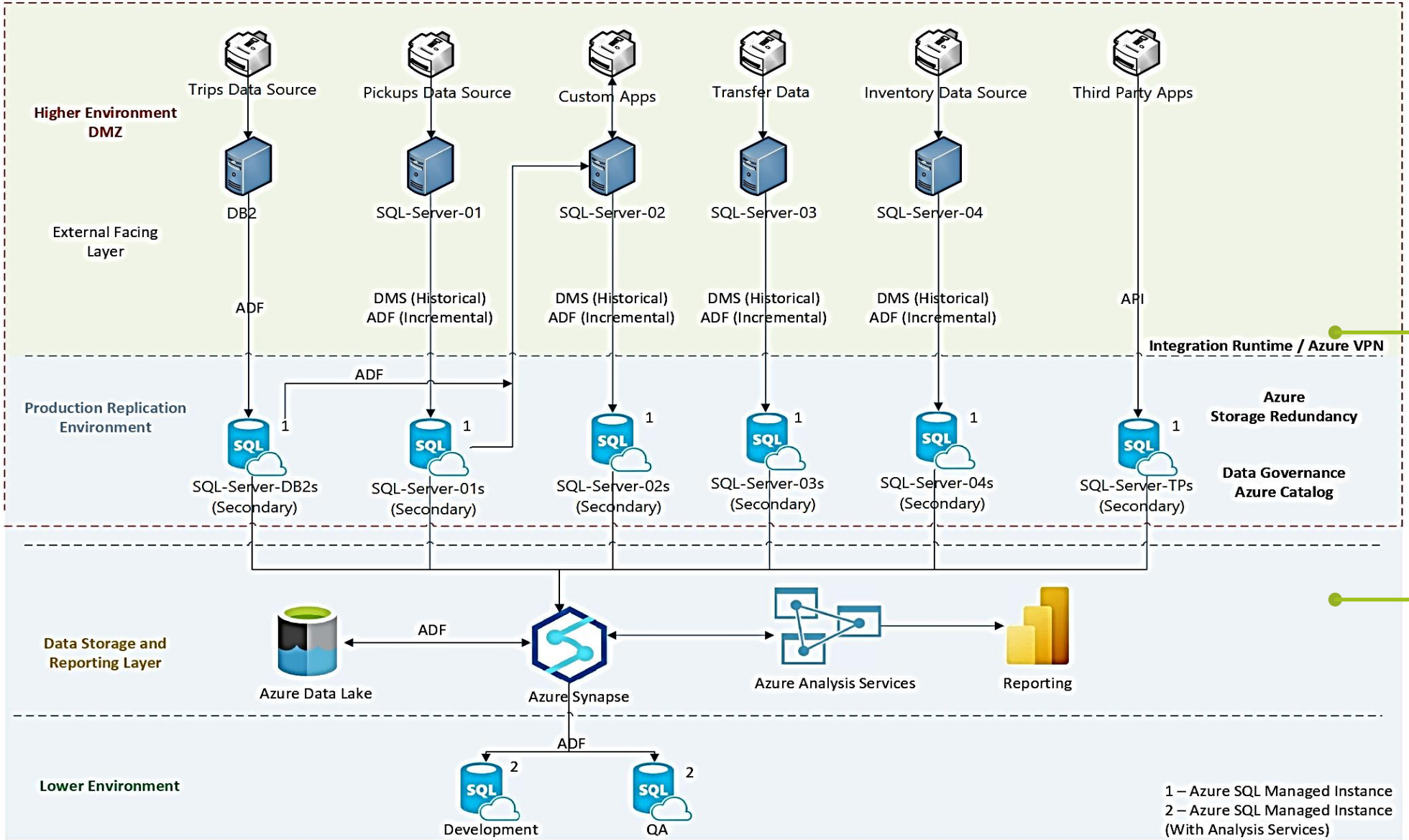
- **Create a DMZ** (Demilitarized Zone) in Higher Environment
- **Replicate data** from primary (source) servers to production replication environment layer (secondary servers)

Establish Online Data Storage (ODS) Layer:

- Push data from secondary servers into a primary ODS server and secondary ODS server
- Use a production reporting server (with SSAS) linked with primary ODS Server to act as a source for reporting
- Add a SAN Storage server to primary ODS server for data archiving purposes

Recommended Cloud-based Network

ILLUSTRATIVE



Connection between On-premises and Azure networks can be established using **Integration Runtime or Azure VPN for migration**

Data storage layer in Cloud:

- Push data from all Azure secondary servers to Azure Synapse which can act as a data storage for all systems
- Use Azure synapse with in-built Analysis Services Integration for reporting into Power BI or Excel