Blueprint™

Sources
Data Remain in-Place

Microsoft SQL Server
MySQL
PostgreSQL
elasticsearch
Azure Data Lake
amazon S3
JSON
Parquet

Data In-Place
Abstracted Access Layer

Optional Architecture. For Consideration:

- Common Security Abstraction
- Enforce & Apply PII/ID Tagging
- Dual-Grant Access Control, MFA Negotiation
- Secrets Mgmt.

CONDUIT

Security & Access Control	Admin & Mgmt. Portal	Monitor Audit Usage	*Data Catalog, Tagging**	*Row/Role Level Security**	*Anonymize & Data Masking**

- SQL Translation + Optimization Engine
- GPU Multi Source Query GPU-Based Engine — Spark

* Roadmap

Data Consumption
Conduit's Common End Point

- Common Data Access Point
- Common Security Access Method, Logged & Audited
- Performance Management
- Intelligent Data Caching

API & Web
tableau
Power BI
jupyter

Sources
For Migration

Files
XML, JSON
RDBMS
API
SaaS / 3Rd Party, Ref.
Streaming

Acquisition
ETL/ELT & Streaming

Integration Patterns

- Append Only
- Historical + Upserts
- Change Capture (CDC)
- Truncate, Repopulate
- Transform, Integrate, Master
- Batch Frequency, 5 min to 24 hr. +
- Metadata & Data Classification

Streaming Data

- IoT/Event Hubs, Kafka
- Captured (AVRO)
- Processed to Lake or Event

Data Lake
Multiprotocol, Schema-Defined & Schema-less, Encrypted

Data Lake

[Staging/..] → [Trusted/..] → [Curated/..]

[Sandbox or User Contributed/..]

Access Layer & Notional Data Segmentation Pattern

Serving DW

Structured, Indexed Store
Policy Mgmt
Data Masking/Encrypted

Alternative Access Layer

Overall Platform Strategy
Engaging for Success

Exploration & Determination

- Primary (MVP) Use Cases & Stakeholder Mgmt., Agile DevOps
- Compliance, Policy, Regulatory
- Regional/Geo Data & Svc Replication
- Data Tiering & Classification, Platform Cost Controls
- SLA & Performance Mgmt.
- Long Term Support Services - Data Source, Virtual Data Set Mgmt.

bpcs.com/conduit