Cloud strategies that work for government
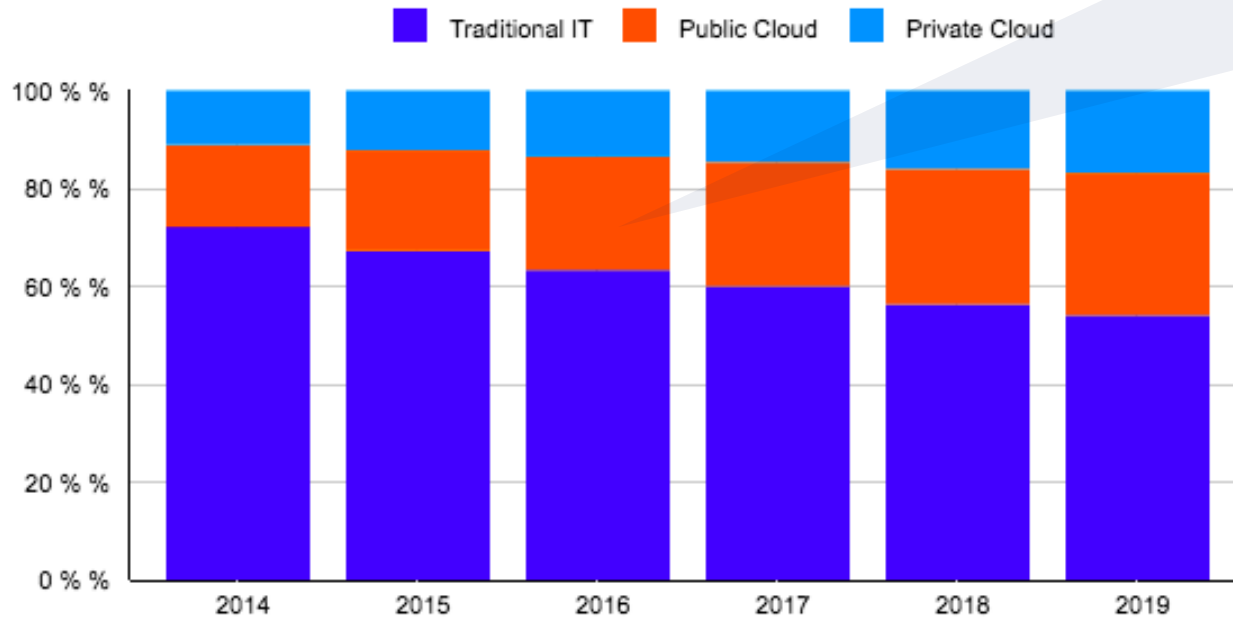
Charles Zedlewski | SVP Products, Cloudera

Public cloud adoption is rising
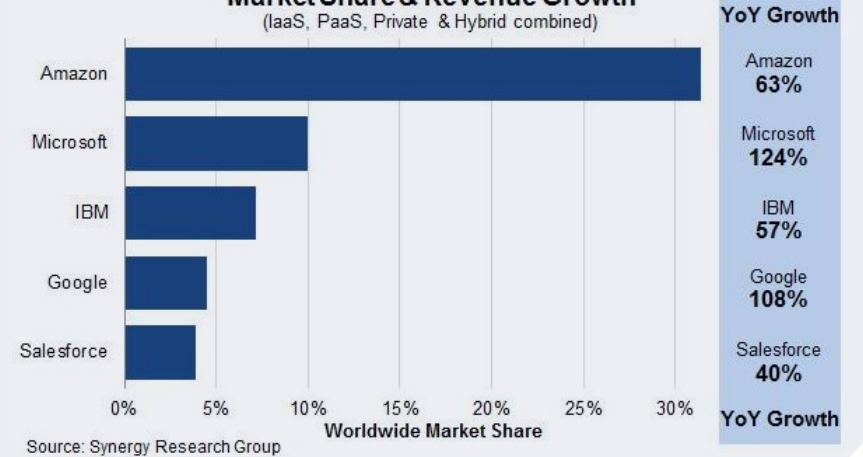
IDC
Analyze the Future

Worldwide Cloud IT Infrastructure Market Forecast by Deployment Type 2014 - 2019 (shares based on Value)

Legend: ■ Traditional IT ■ Public Cloud ■ Private Cloud

Source : IDC Worldwide Quarterly Cloud IT Infrastructure Tracker Q2 2015

Cloud Infrastructure Services - Q4 2015
Market Share & Revenue Growth
(IaaS, PaaS, Private & Hybrid combined)

	YoY Growth
Amazon	63%
Microsoft	124%
IBM	57%
Google	108%
Salesforce	40%

Worldwide Market Share
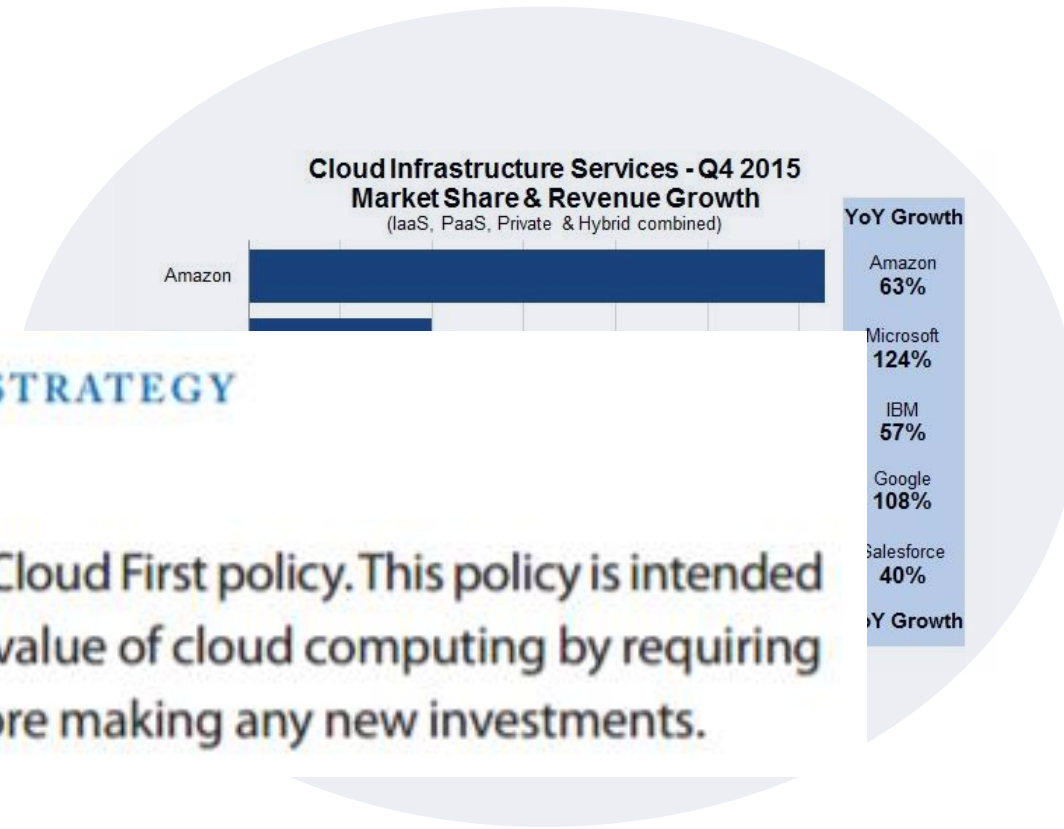
Source: Synergy Research Group

Public cloud adoption is rising

FEDERAL CLOUD COMPUTING STRATEGY

To harness the benefits of cloud computing, we have instituted a Cloud First policy. This policy is intended to accelerate the pace at which the government will realize the value of cloud computing by requiring agencies to evaluate safe, secure cloud computing options before making any new investments.

Cloud Infrastructure Services - Q4 2015
Market Share & Revenue Growth
(IaaS, PaaS, Private & Hybrid combined)

	YoY Growth
Amazon	**63%**
Microsoft	**124%**
IBM	**57%**
Google	**108%**
Salesforce	**40%**

Cloudera customers are leading the way

One platform supports 4 kinds of applications

Data Engineering

Data Science

Analytic DBMS

Operational DBMS

ANALYZE

BATCH Spark, Hive, Pig MapReduce	STREAM Spark	SQL Impala	SEARCH Solr	OTHER Kite

UNIFIED SERVICES

RESOURCE MANAGEMENT YARN	SECURITY Sentry, RecordService

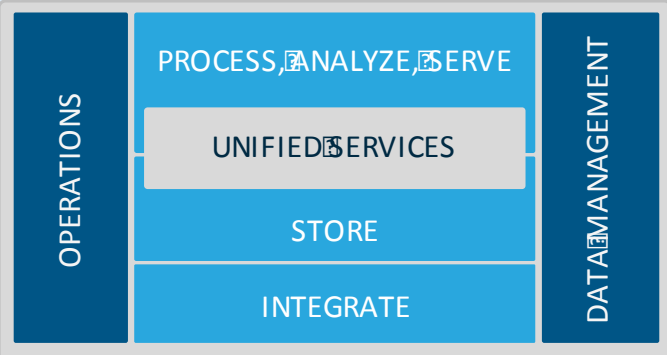| FILESYSTEM HDFS	RELATIONAL Kudu	NoSQL HBase	OTHER Object Store

STORE

OPERATIONS
Cloudera Manager
Cloudera Director

DATA MANAGEMENT
Cloudera Navigator
Encrypt and KeyTrustee
Optimizer

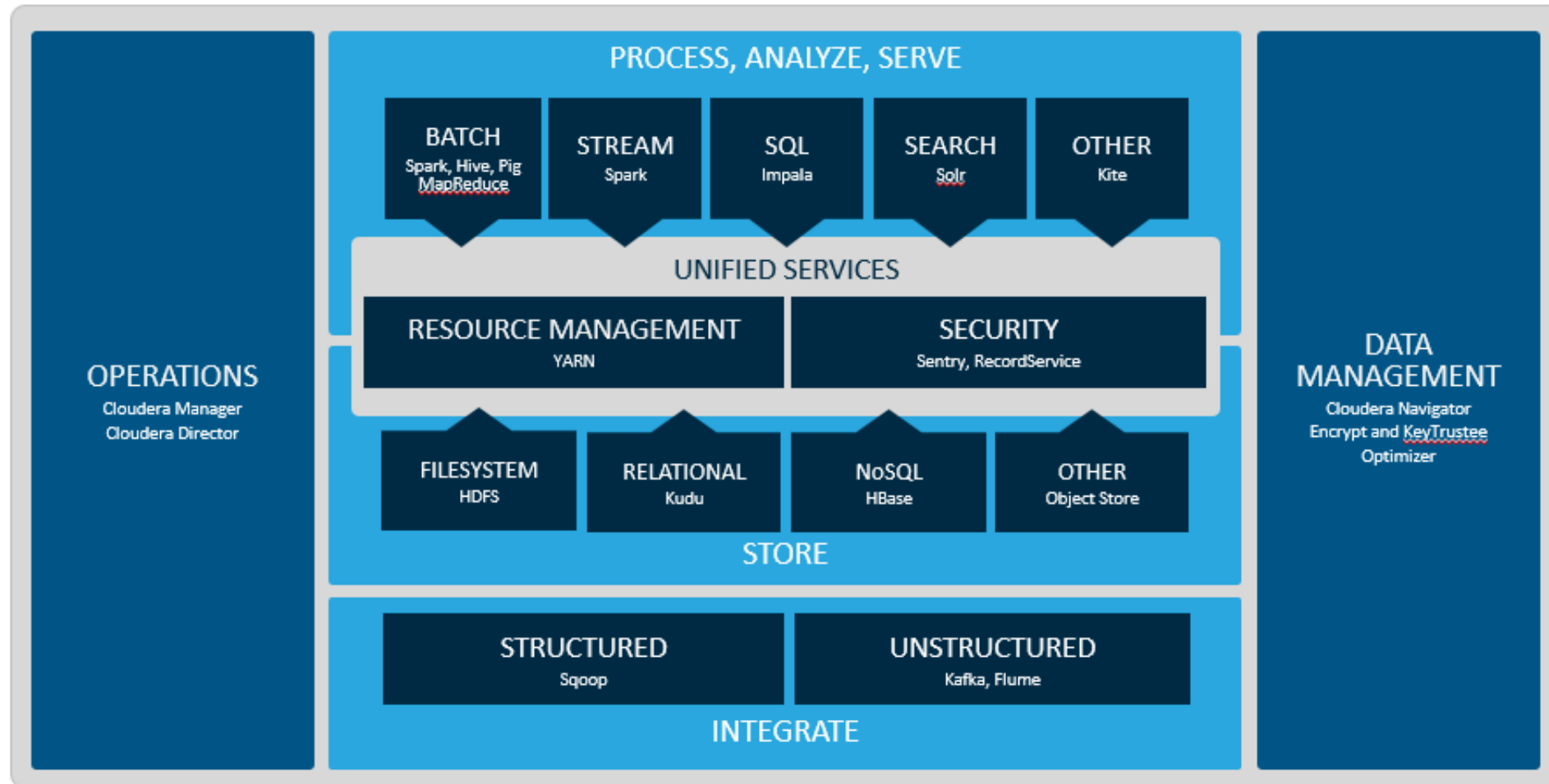| STRUCTURED Sqoop	UNSTRUCTURED Kafka, Flume

INTEGRATE

Really, different configurations of the same platform

Data Science & Engineering
Store and process unlimited data fast and cost-effectively.

Analytic Database
Explore, analyze, and understand all your data.

Real-Time Applications
Build data-driven products to deliver real-time insights.

OPERATIONS

PROCESS, ANALYZE, SERVE

UNIFIED SERVICES

STORE

INTEGRATE

DATA MANAGEMENT

In many environments multiple applications share a single, multi-tenant cluster

In cloud environments, often more separate, specialized clusters tuned for each application

Object Store

Object Store

HDFS

amazon web services

Microsoft
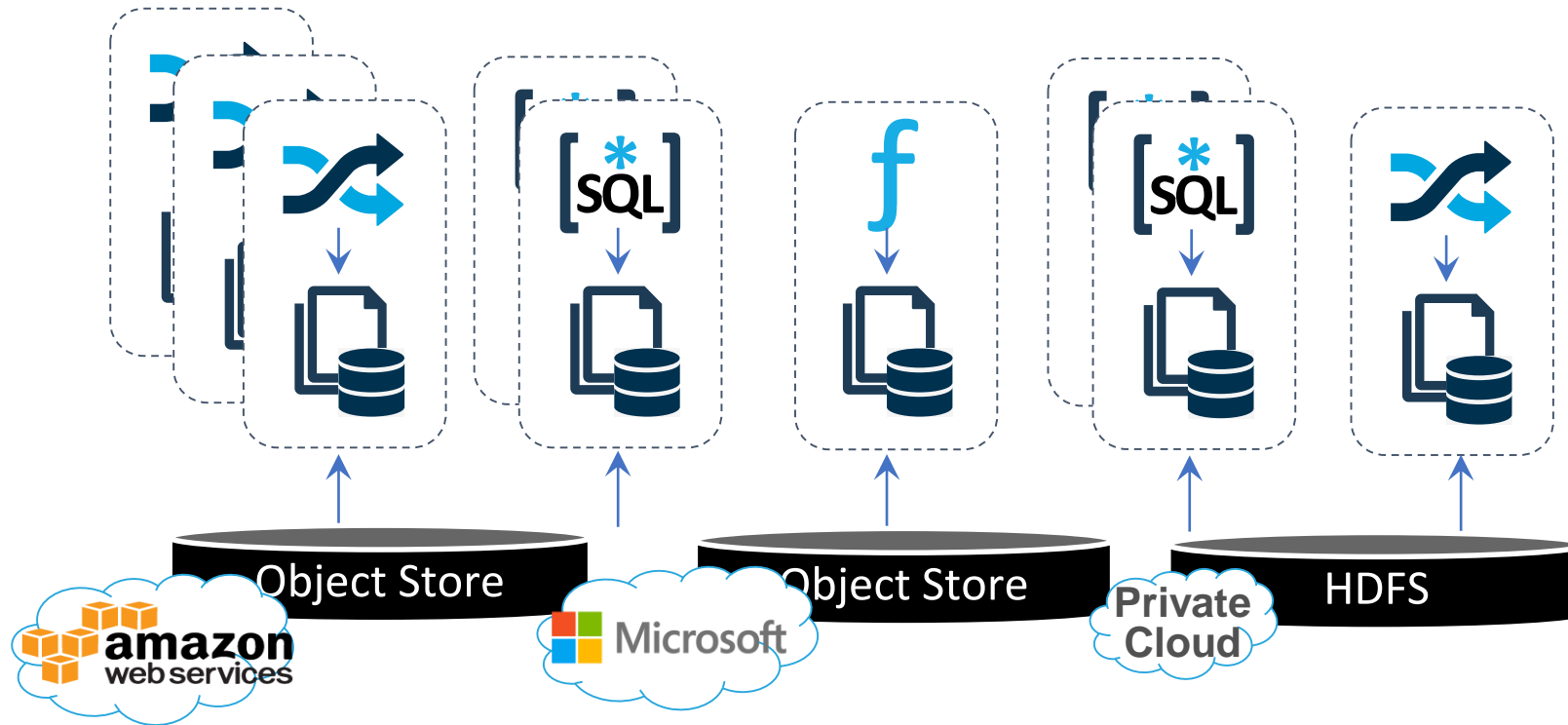
Private Cloud

But openness is even more important in the cloud

Open Environment

Run the same platform in different clouds or on bare metal, so customers can move as needed without migration or retraining

Open Ecosystem

1000's of applications that need to run on the platform with the assurance of compatibility across releases and clouds

Open Source

Avoid vendor lock-in, and leverage components supported by the committers who drive the community roadmap

But openness is even more important in the cloud

Open

Open

Open

Be Open First

In supporting our mission, one of our goals is to "Be Open" by enabling the American people and an increasingly mobile workforce to access high-quality Agency data, APIs, and code anywhere, anytime, on any device. Thus, unlocking the power of government data to spur innovation across our nation and improve the quality of services for the American people.
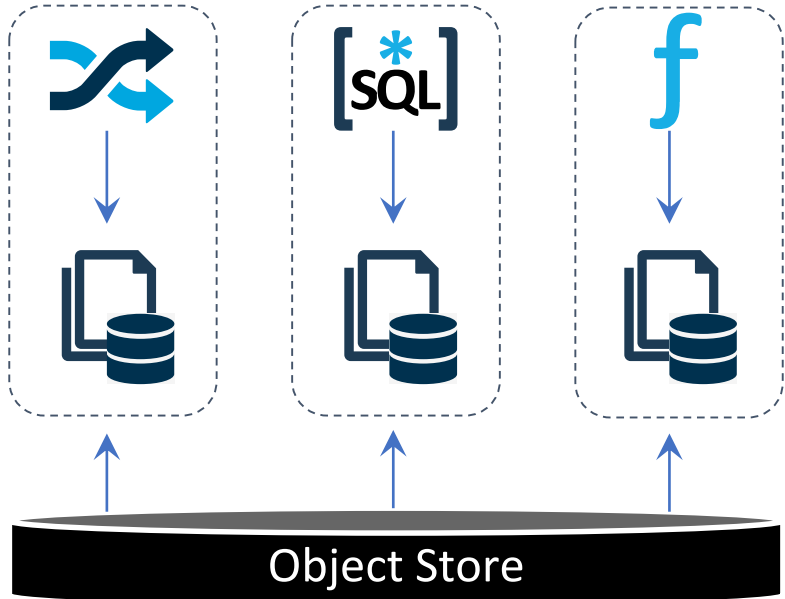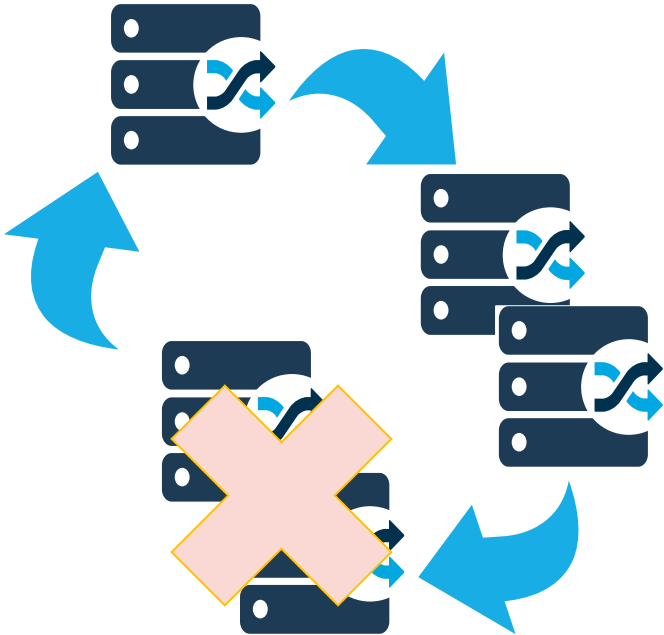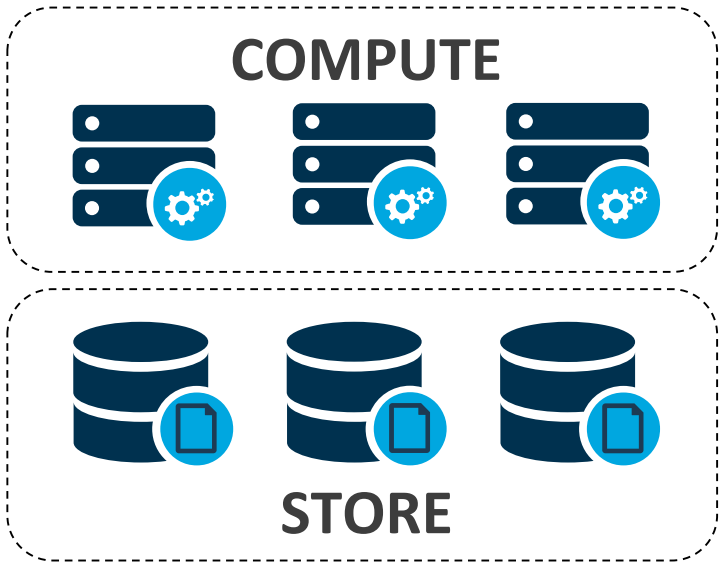
Optimizing Cloudera for cost & convenience in the cloud

cloudera

Optimizing Data Engineering & Data Science workloads for the cloud

Hybrid support for multiple environments

Transience for flexibility, lower TCO and risk

Unified platform, from ingest to insight and action

COMPUTE

STORE

[SQL]

Object Store

Benefits of Data Engineering with Cloudera
Lower TCO and increased flexibility on a trusted enterprise data platform

Lower TCO	Increased Flexibility	End-to-End for the Enterprise

Manage Costs
- Transience for dev/test, ETL, and data science
- Usage-based pricing
- Spot instance support

Multi-cloud
- Shop across providers: Amazon, Google, Microsoft

Deliver On-Demand
- Immediate access to large compute with fast cluster provisioning
- Self-service for developers

Optimize and Isolate
- Tailor infrastructure for the job
- Run different software versions
- Enable more experimentation with less opportunity cost

Build Complete Data Apps
- Ingest, stream, process, explore analyze, model, and serve on the same platform
- Shared data with object store integration
- Cluster metadata persistence
- Common compliance-ready security and governance frameworks

Optimizing Analytic DBMS workloads for the cloud

Reduce Operating Costs

ETL

Only pay for what you need, when you need it

- Transient clusters
- Object storage centric
- Cloud-native deployment
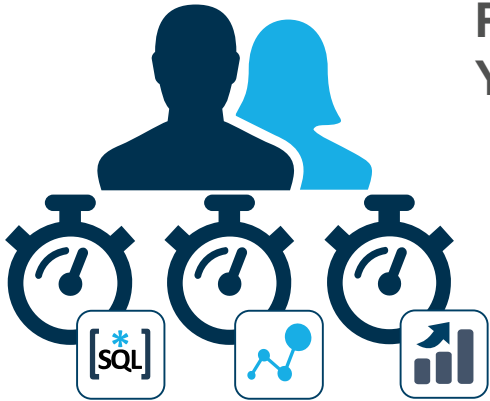
New Insights, New Revenue

BI/Analytics

Explore and analyze all data, wherever it lives

- Long-running clusters
- Object storage or local storage
- Lift-and-shift deployment

Benefits of Cloudera's Analytic Database
ETL and BI/Analytics in the cloud

Predictable Results Whenever You Want
- Consistent query performance, even during peak times
- Multi-tenancy via isolated clusters on shared data

Add Use Cases, Analytics, and Data On-Demand
- Avoid the IT backlog with instant access to all data
- On-demand clusters query directly on shared object storage

Contention-Free ETL
- ETL anytime without impacting other workloads or risking SLAs
- Separate ETL clusters as-needed on shared data

Just-in-Time Resources
- Real-time capacity for your needs, as they change
- Elastically grow/shrink your cluster via decoupled architecture

Operational Database

Web-Scale Data Depot

Durable, low latency storage for web applications, message stores, and mission critical operational activities.

Complex Event Processing

Identifying meaningful events based on multiple data streams and taking action.

Model Scoring/Serving

Use data and current/past events to score and serve the likelihood of subsequent events.

Operational Database in the Cloud

Public Cloud Benefits

Cost Goals

Convenience Goals

- Low-cost backup and disaster recovery
- Development and testing environments easy to deploy and decommission

- Elastic growth for tightly provisioned workloads makes expansion easy, and enables a lower-cost steady state

- Fast and easy provisioning of additional clusters helps projects move quickly

Pulling it together

cloudera

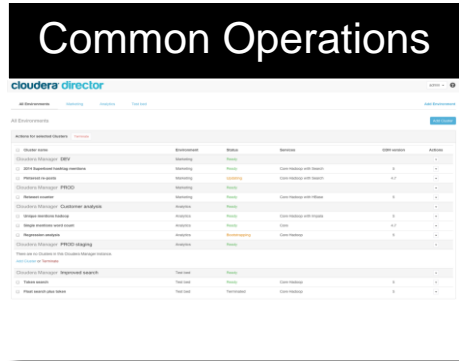# The cloud creates more & smaller specialized clusters for each application—which can turn into silos
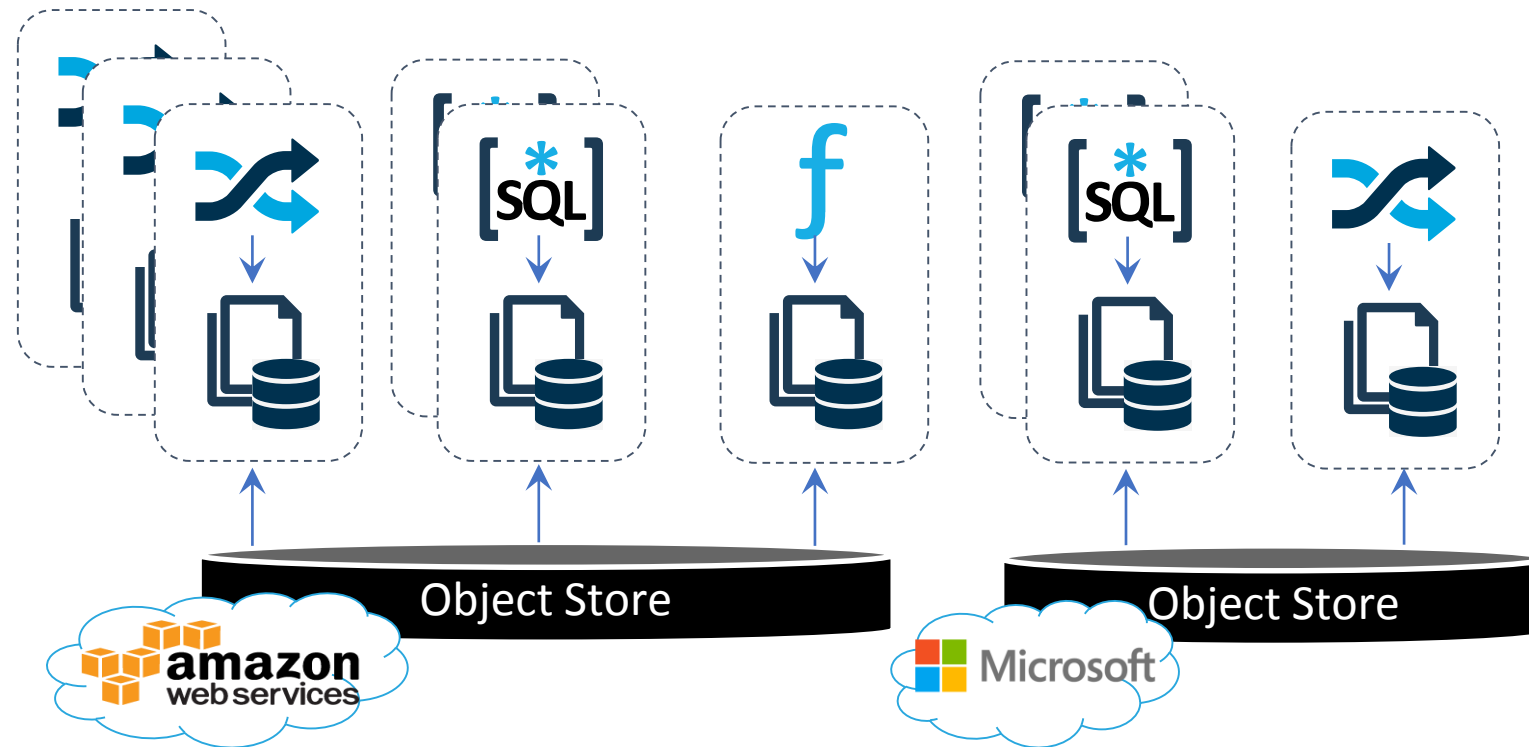
And many problems are a combination of SQL & predictive, batch & online
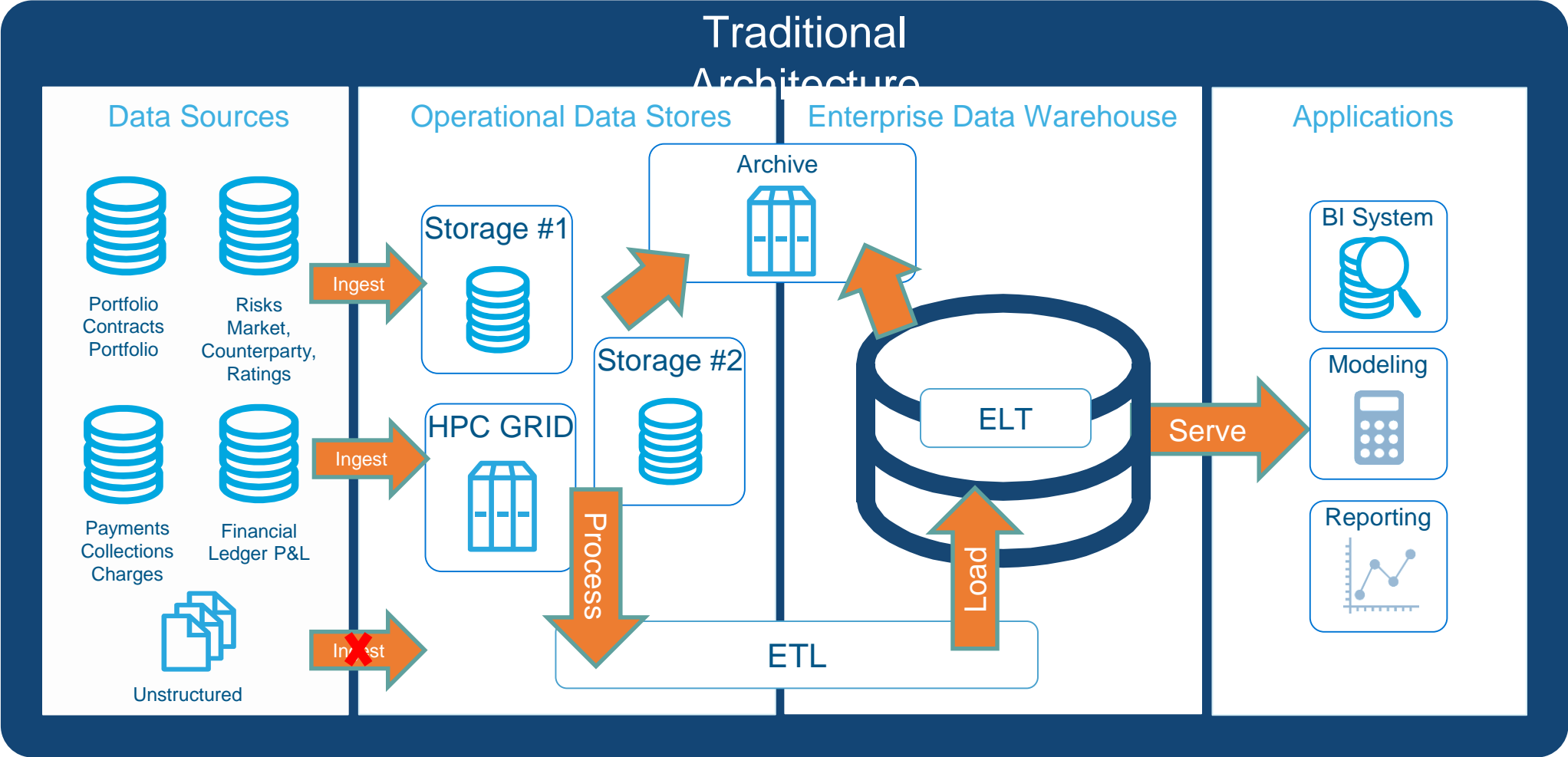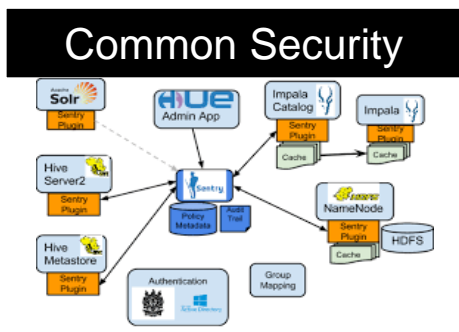
An Enterprise Data Hub in the cloud
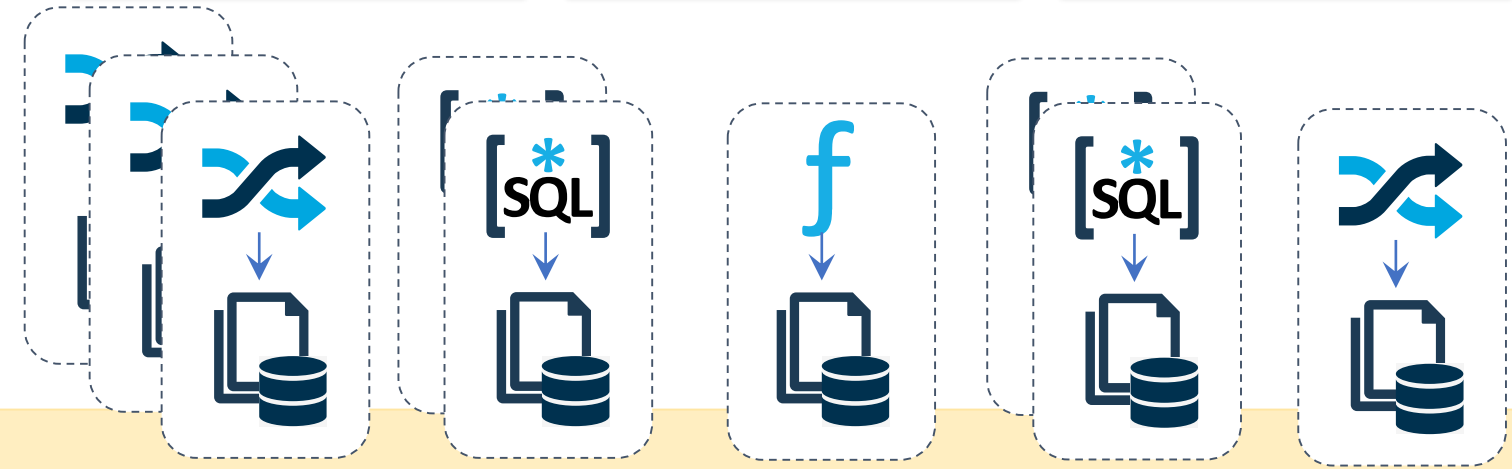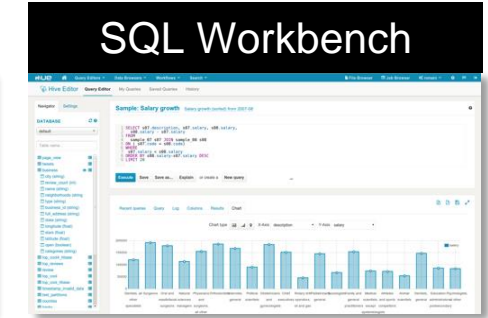

Common Operations

Common

Common Security

Developer

Partner Ecosystem

SQL Workbench

Common: Operations, Governance, Security, Schema, Catalog

Object Store

Object Store

amazon web services

Microsoft

Thank you