March 4, 2020

Cloud Data Warehouse Migration for PowerCenter Customers

Ron Lunasin, Sr. Director
Strategic Ecosystems and Solution Architecture

Informatica

Journey to Cloud Data Warehouses and Data Lakes

Customer Use Cases	Data Warehouse As-Is Lift and Shift	New Cloud Data Warehouse	Cloud Data Warehouse Migration	Cloud Data Lake Modernization
	Save on data center costs	Start small and fast and grow as needed	Adopt cloud analytics gradually over time	End to end modern data infrastructure for next gen analytics
Technology Examples	teradata. → teradata. ORACLE EXADATA → ORACLE EXADATA	amazon REDSHIFT snowflake SQL	teradata. → snowflake NETEZZA → amazon REDSHIFT ORACLE EXADATA → SQL	CLOUDERA → Azure Data lake MAPR → aws databricks DELTA LAKE

Informatica

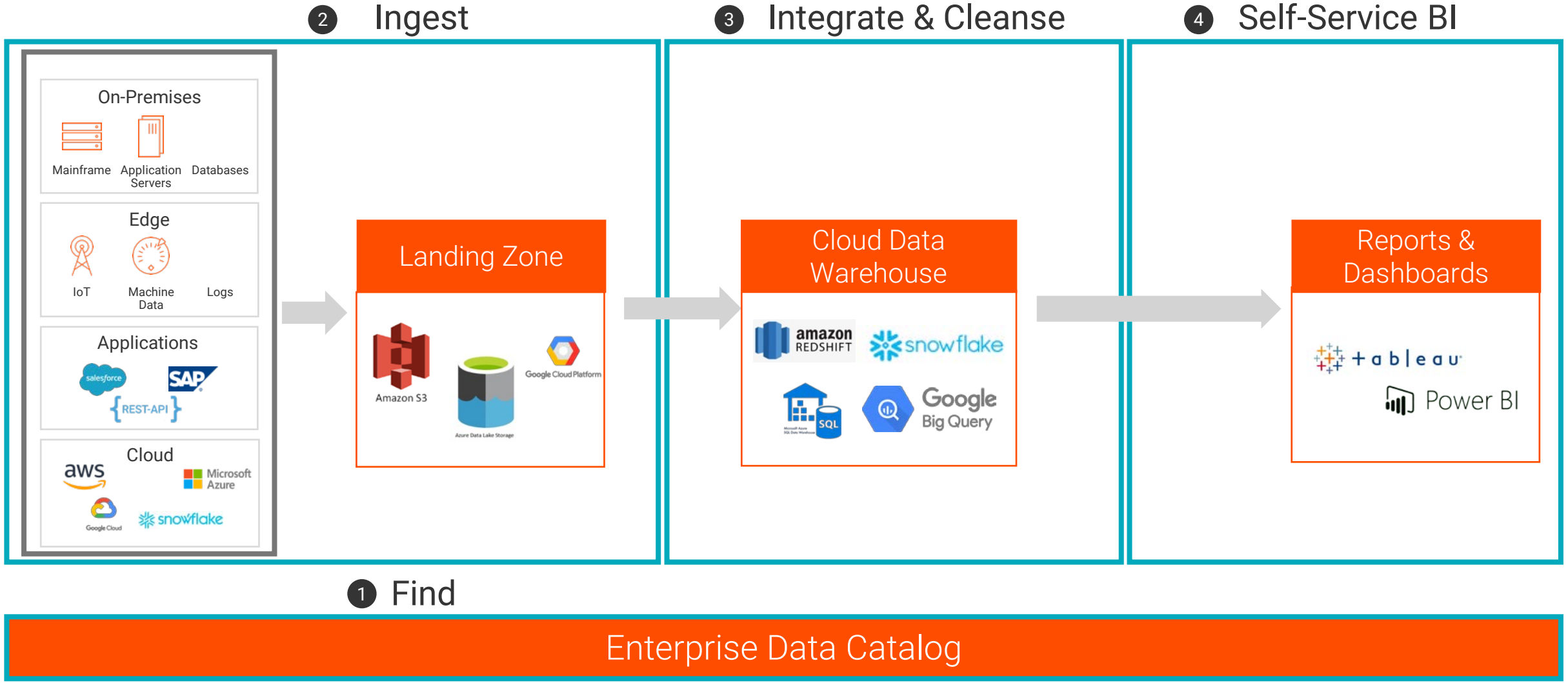
Goals:

- Focused BI, Departmental reporting or Self-Service Analytics initiative - LoB or project
- Agility
- Time to value

Challenges:

- Hand-coding is difficult to maintain
- Difficult to on-board new data sources
- Finding and getting data from existing on-prem sources
- Poor data quality

New Cloud Data Warehouse



Informatica

Goals:

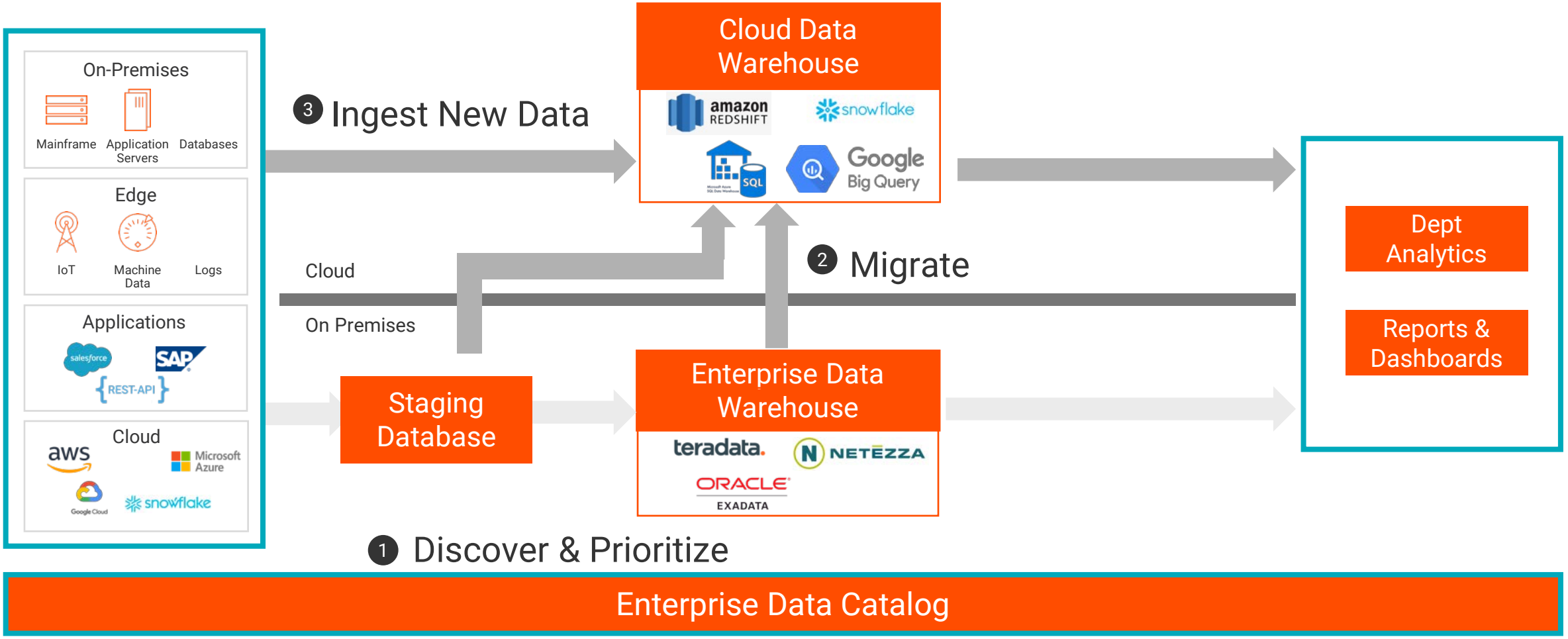
- Cloud First Strategy. Adopt Cloud EDW on PaaS (AWS / Azure / GCP)
- Adjust analytics and integration to the new CDW
- Don't want to propagate bad quality data into the new CDW
- Rapidly adopt new innovations in the cloud

Challenges:

- Too difficult to know and prioritize which data to migrate to new CDW
- Can't leverage existing investments in ETL/ELT; must rebuild everything from scratch
- Poor data quality

Cloud Data Warehouse Migration

Solve the immediate problem/need and position for the future



Informatica

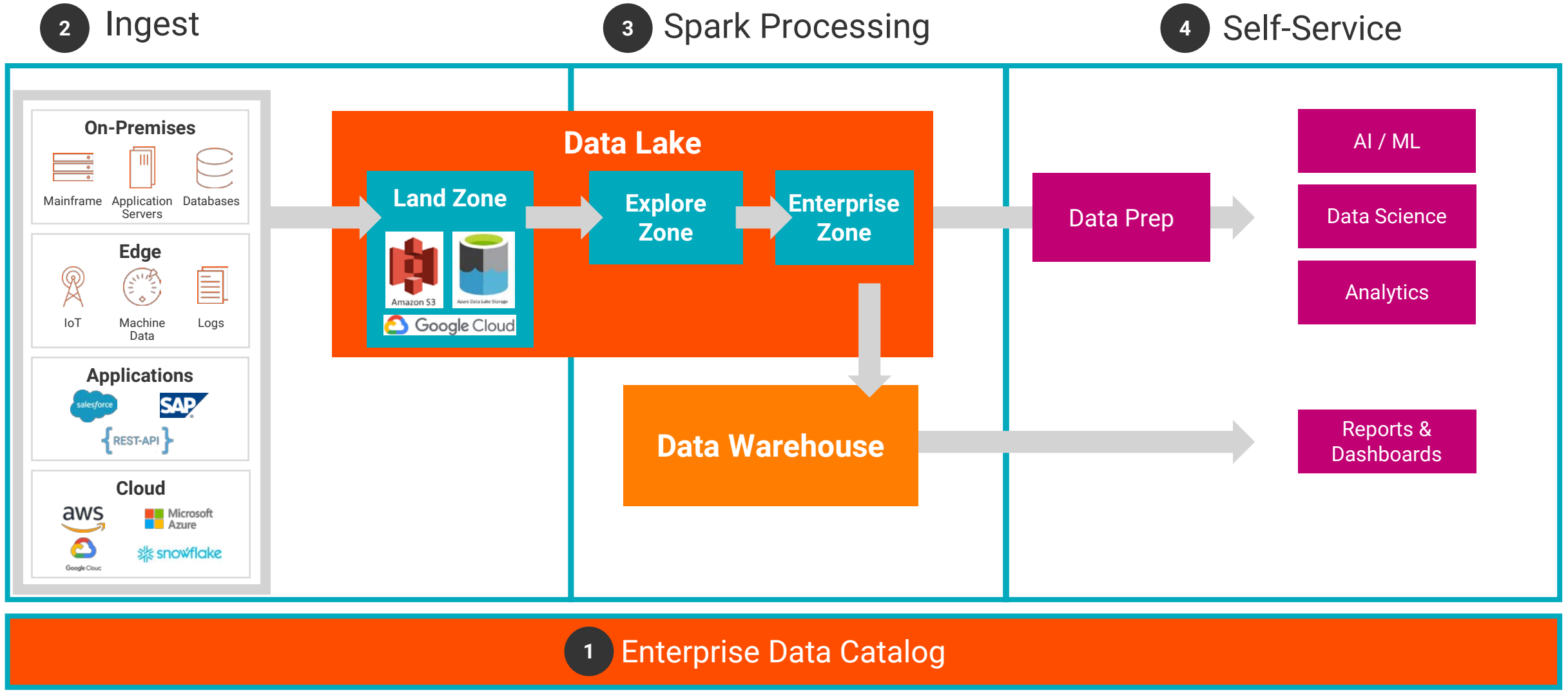
Goals:

- Data platform that can go beyond traditional analytics:
  - AI/ML
  - Data Science
  - Self-Service Analytics
  - Real-Time Analytics
- Performance at massive scale
- New processing engines (e.g. Spark, Databricks)
- Serverless and elastic
- New types of emerging data sources such as machine, social, etc.

Challenges:

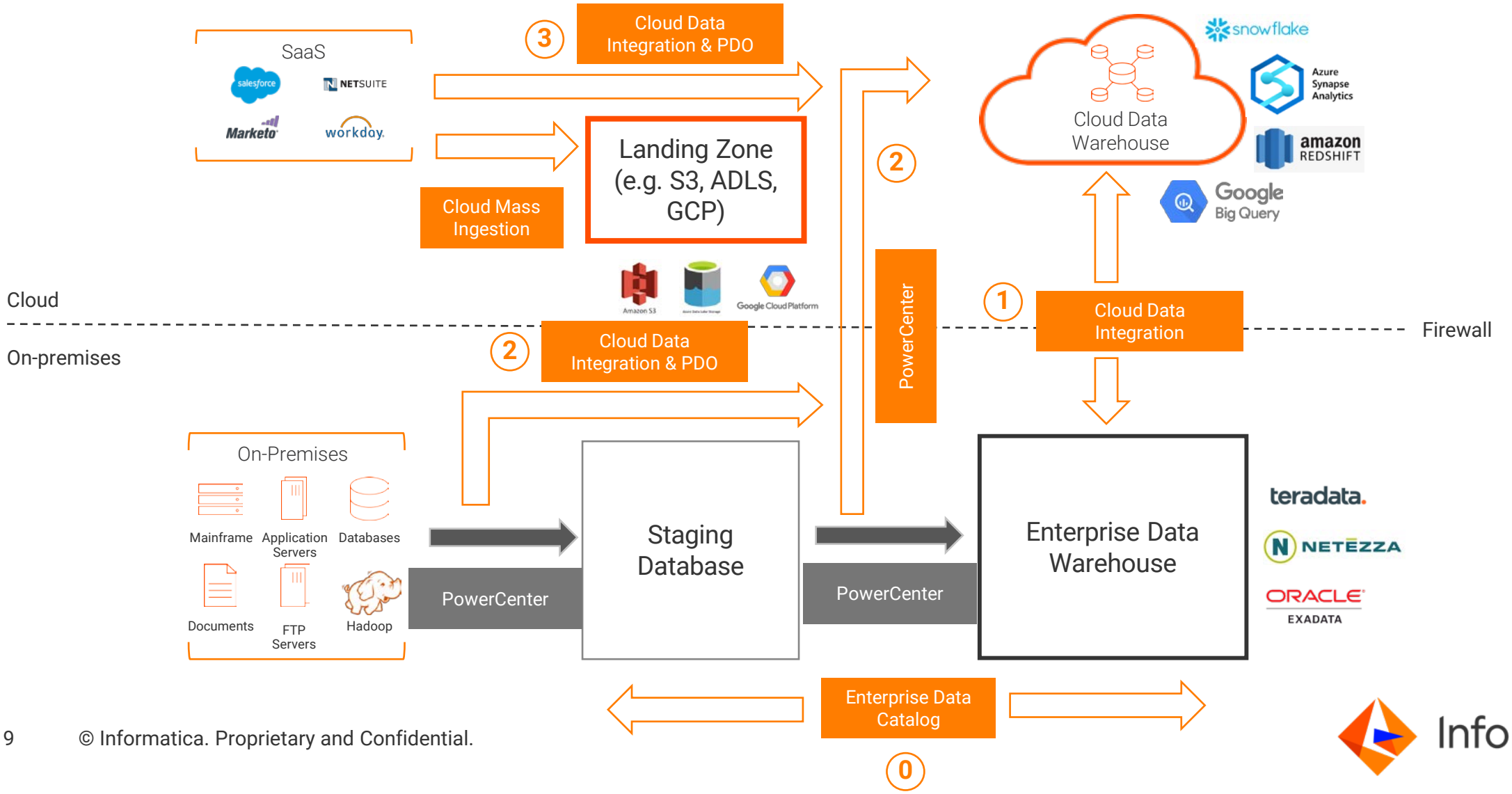
- Keeping up with a rapidly changing technology landscape and ecosystem
- Too difficult to find and use the right data
- Takes too long to get data pipelines into production and get value out of analytic projects

Cloud Data Lake architecture



Informatica

Agile Phased Migration



Informatica

To discuss best practices contact icare@informatica.com

Demo: Quickly find and prioritize data to migrate using the Enterprise Data Catalog

Q & A

Informatica

Poll Question

Which of the following traditional data warehouses do you currently use (check all that apply)?

A. Teradata
B. Netezza
C. Oracle
D. Microsoft
E. DB2
F. Other

Informatica

Poll Question

Which of the following cloud data warehouses do you plan to use (check all that apply)?

A. Snowflake
B. AWS Redshift
C. Azure Synapse
D. Google BigQuery
E. Teradata Vantage
F. Other