The open source serverless data lakehouse - the next evolution in big data

romeo kienzler

19th of Oct 2022, for Baselone.ch

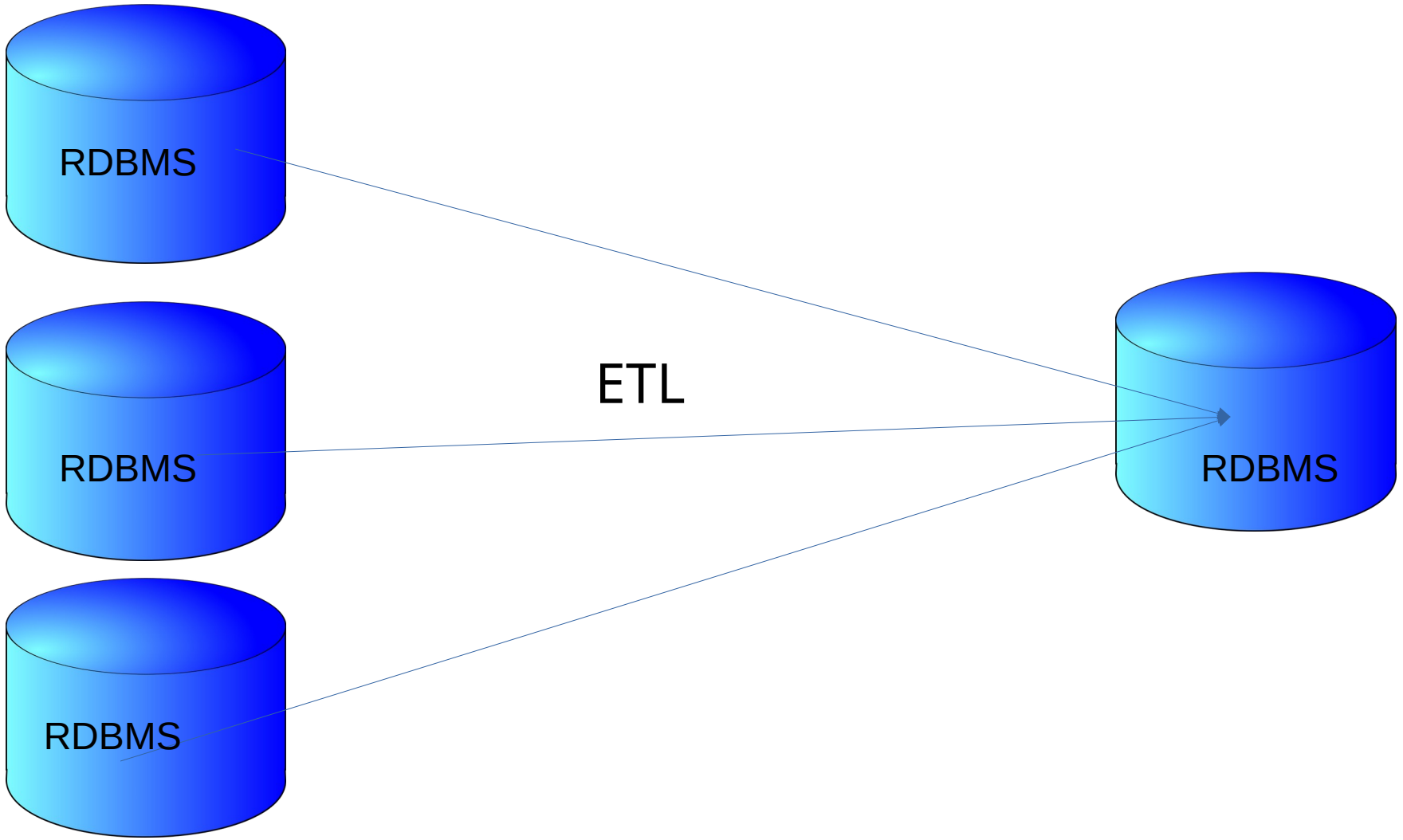
CREDITS

Torsten Steinbach (torsten@de.ibm.com) – CTO Big Data in Cloud, IBM

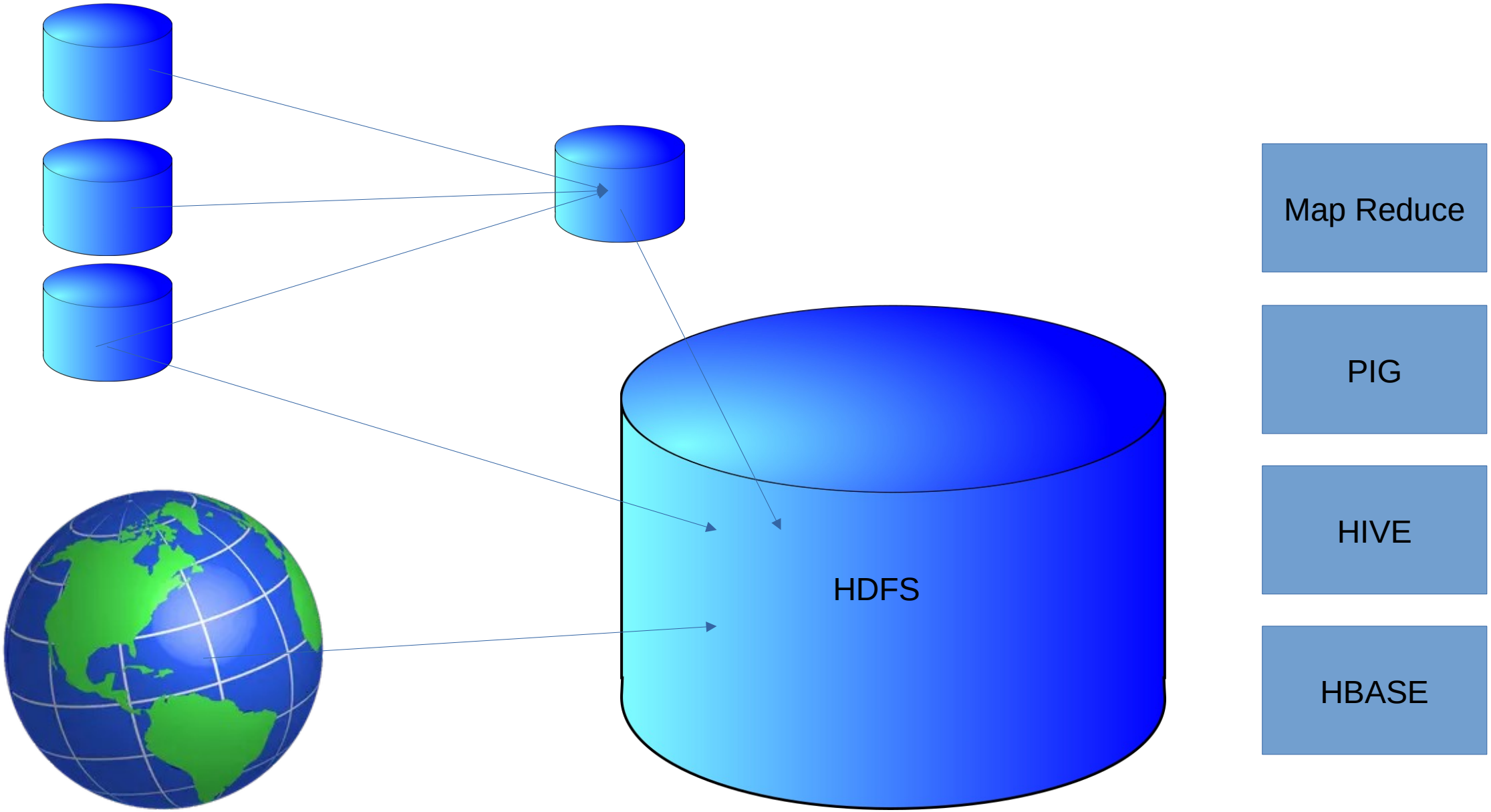
https://community.ibm.com/community/user/cloud/viewdocument/modernize-your-big-data-analytics-w
https://ibm.co/3ySkeTI

July 2022

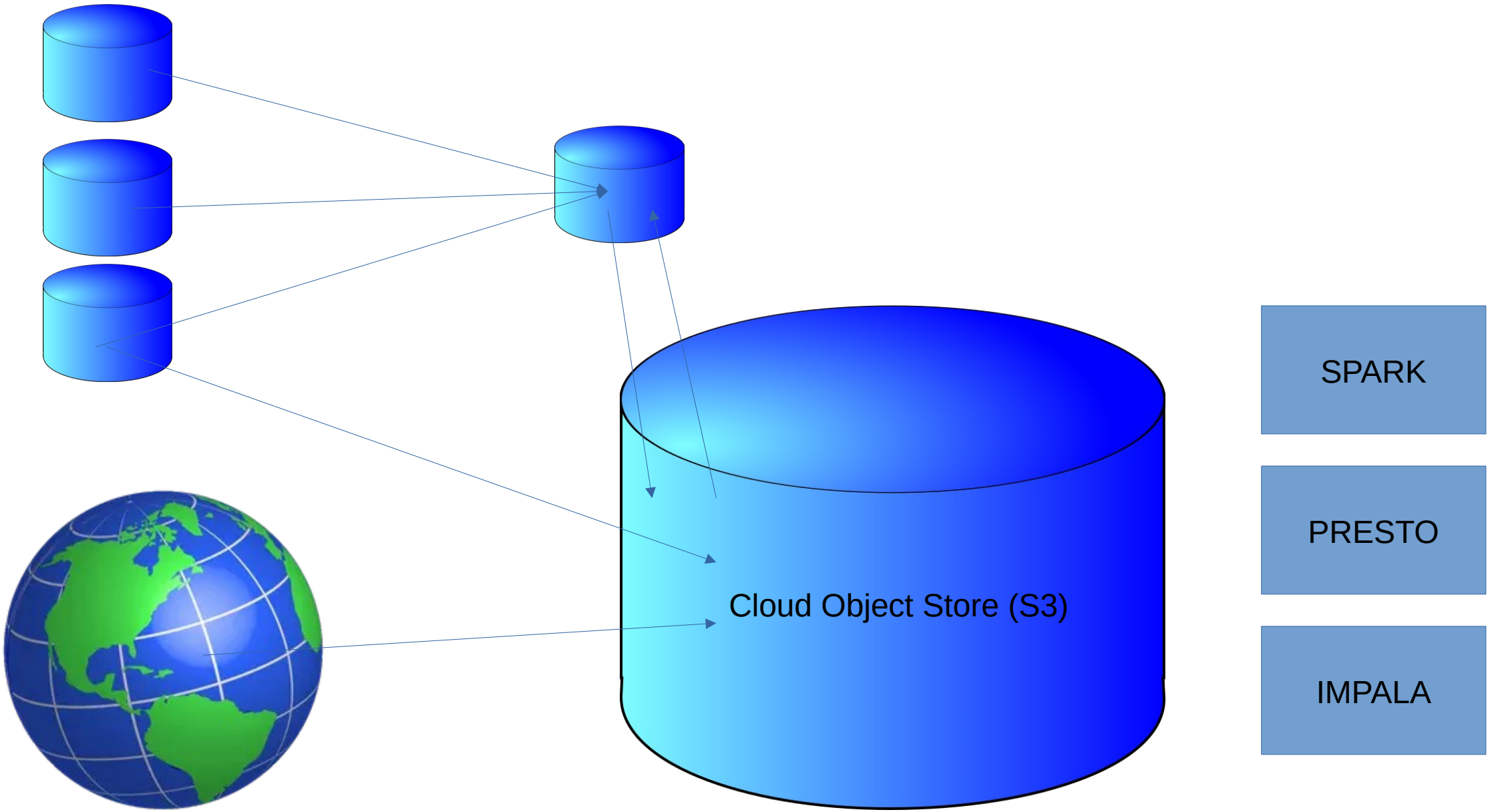
Enterprise Data Warehouses (Level 1)



Hadoop Data Lakes (Level 2)



Cloud Data Lakes (Level 3)



Stocator: Providing High Performance and Fault Tolerance for Apache Spark Over Object Storage

Publisher: IEEE

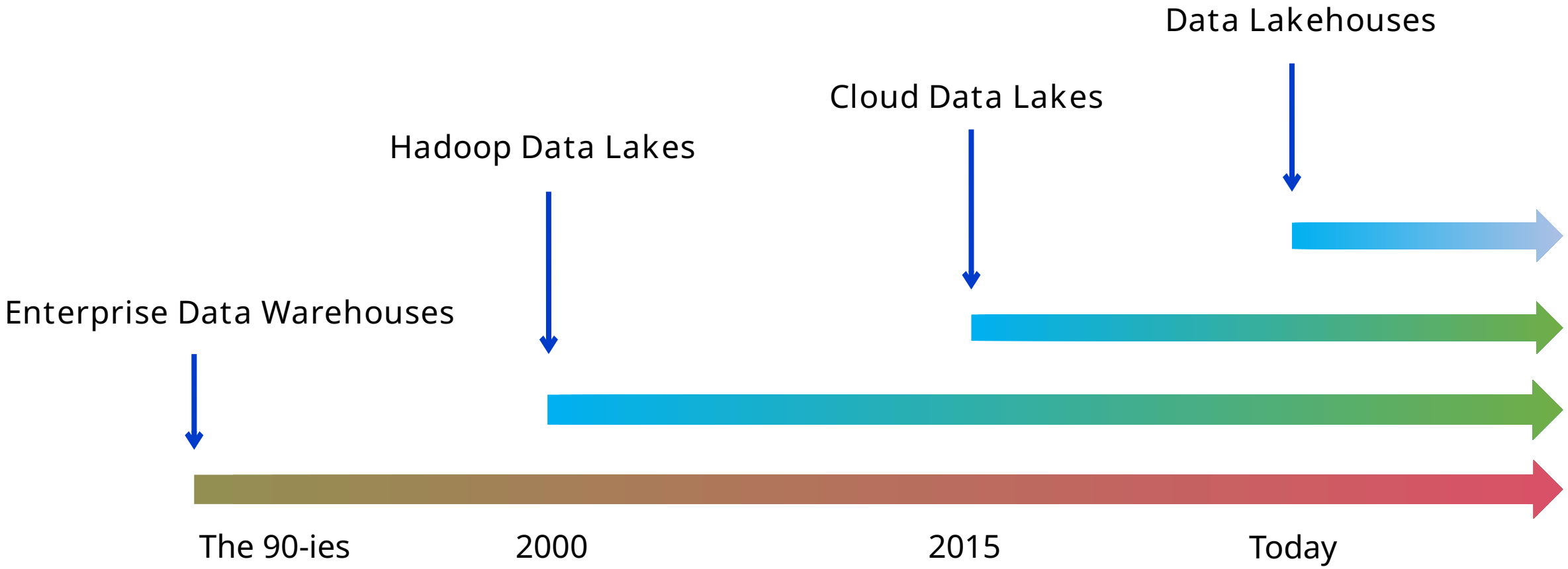
Gil Vernik ; Michael Factor ; Elliot K. Kolodner ; Pietro Michiardi ; Effi Ofer ; Francesco Pace **All Authors**

Abstract:

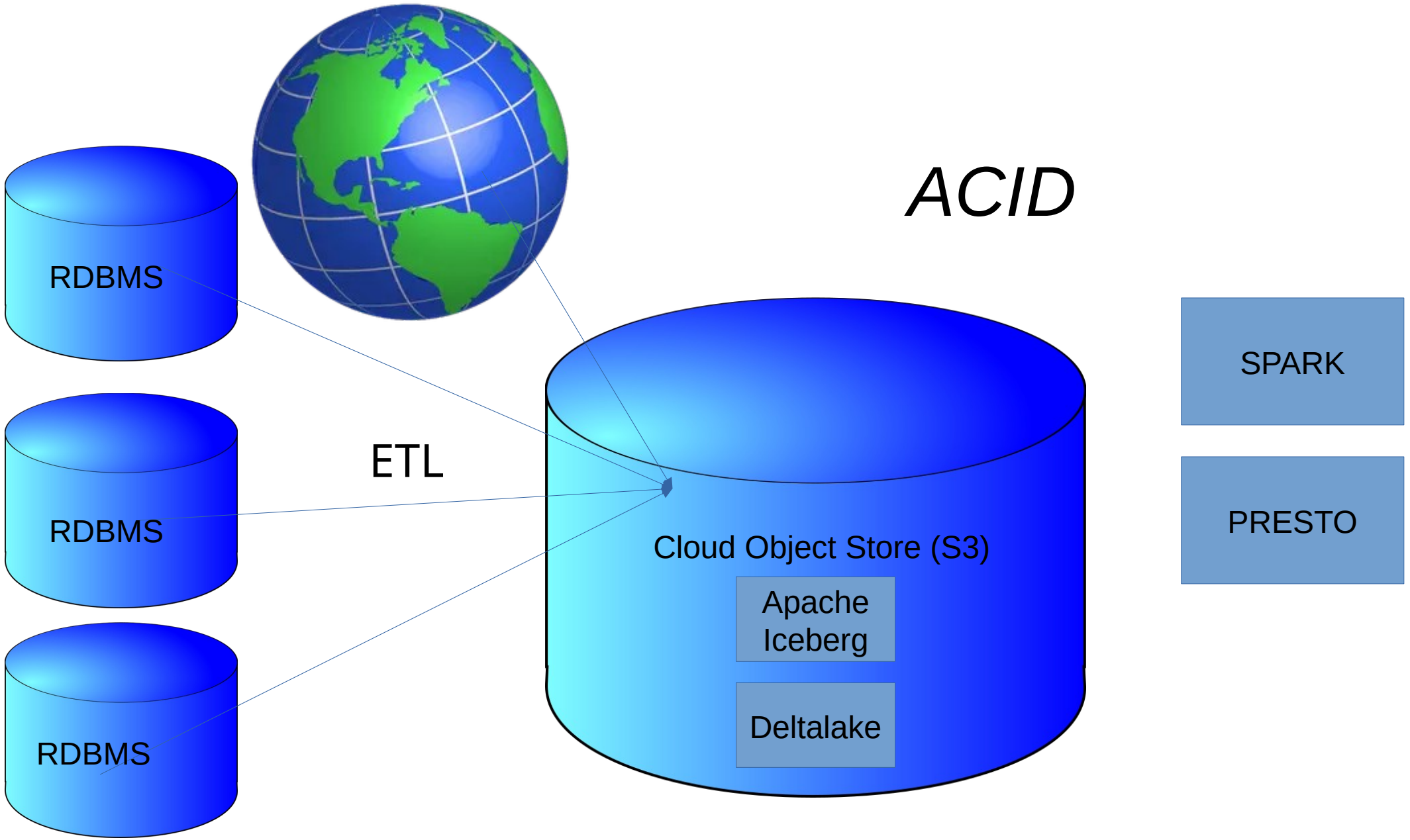
Until now object storage has not been a first-class citizen of the Apache Hadoop ecosystem including Apache Spark. Hadoop connectors to object storage have been based on file semantics, an impedance mismatch, which leads to low performance and the need for an additional consistent storage system to achieve fault tolerance. In particular, Hadoop depends on its underlying storage system and its associated connector for fault tolerance and allowing speculative execution. However, these characteristics are obtained through file operations that are not native for object storage, and are both costly and not atomic. As a result these connectors are not efficient and more importantly they cannot help with fault tolerance for object storage. We introduce Stocator, whose novel algorithm achieves both high performance and fault tolerance by taking advantage of object storage semantics. This greatly decreases the number of operations on object storage as well as enabling a much simpler approach to dealing with the eventually consistent semantics typical of object storage. We have implemented Stocator and shared it in open source. Performance testing with Apache Spark shows that it can be 18 times faster for write intensive workloads and can perform 30 times fewer operations on object storage than the legacy Hadoop connectors, reducing costs both for the client and the object storage service provider.

Published in: 2018 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)

Evolution of Big Data Systems



Data Lakehouses (Level 4)



Technology support

Storage

- Cloud Object Store
  - IBM Cloud Object Storage, Amazon S3, Azure Blob Store, Google Cloud Storage
- On-prem
  - CEPH, MINIO
  - HDFS, IBM GPFS

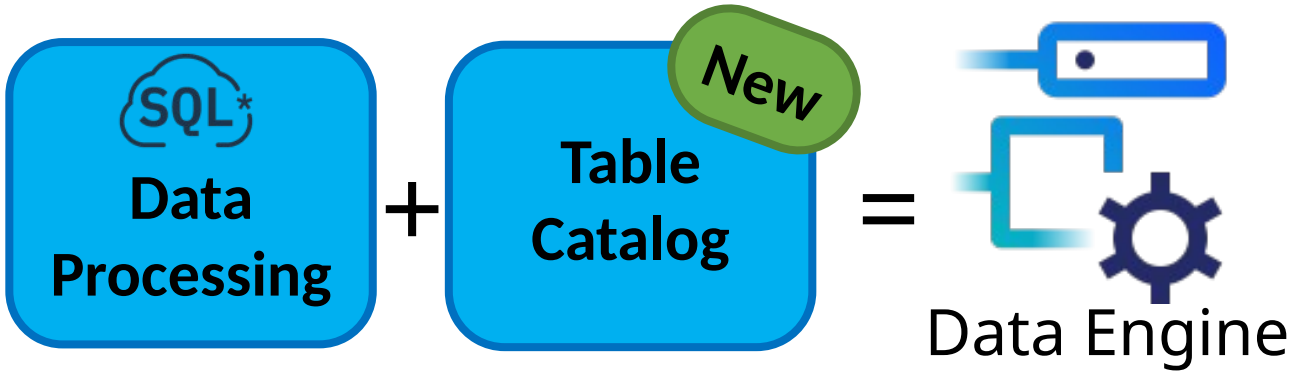
Catalog

- HIVE Metastore
- Nessie

Table Container

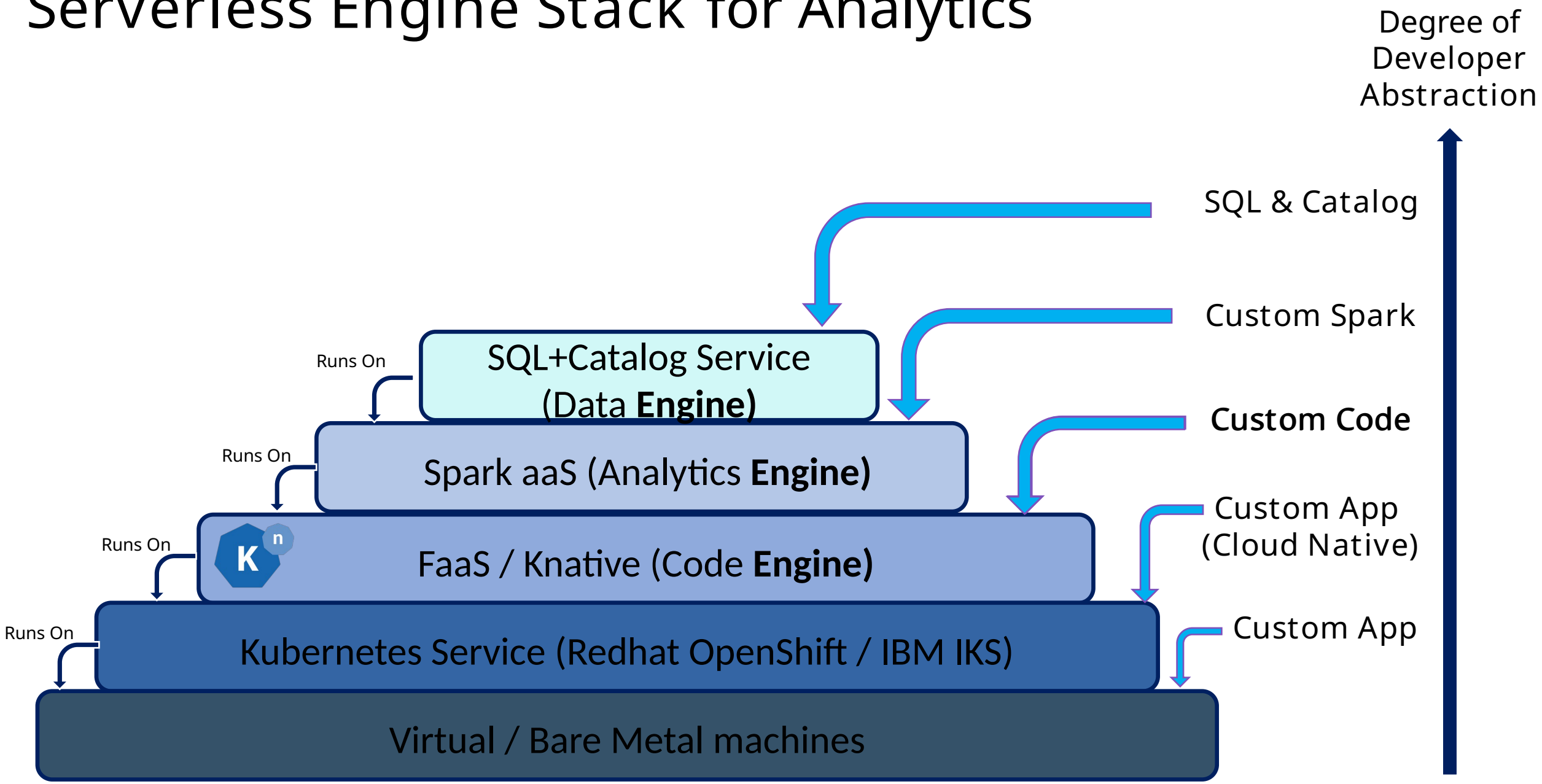
- Hive, ORC, Deltalake, Apache Iceberg

IBM Cloud Data Engine (Level 3)



Aka: SparkSQL + Hive Metastore aaS (Serverless)

Serverless Engine Stack for Analytics



Future work:

Level 3 + ACID + real-time queries = Level 4

Thank YOU!

https://github.com/IBM/claimed

https://github.com/IBM/claimed/blob/master/component-library/transform/spark-sql-interactive.ipynb