



.....

Modern Data Warehouse Platforms & Architectures for the Cloud

June 2022

A PDF of these four charts can be downloaded at
<https://wintercorp.com/tdwi-course-addendum>

Richard Winter, CEO

Dr. Norbert Kremer, Director,
WinterCorp Cloud Database Lab

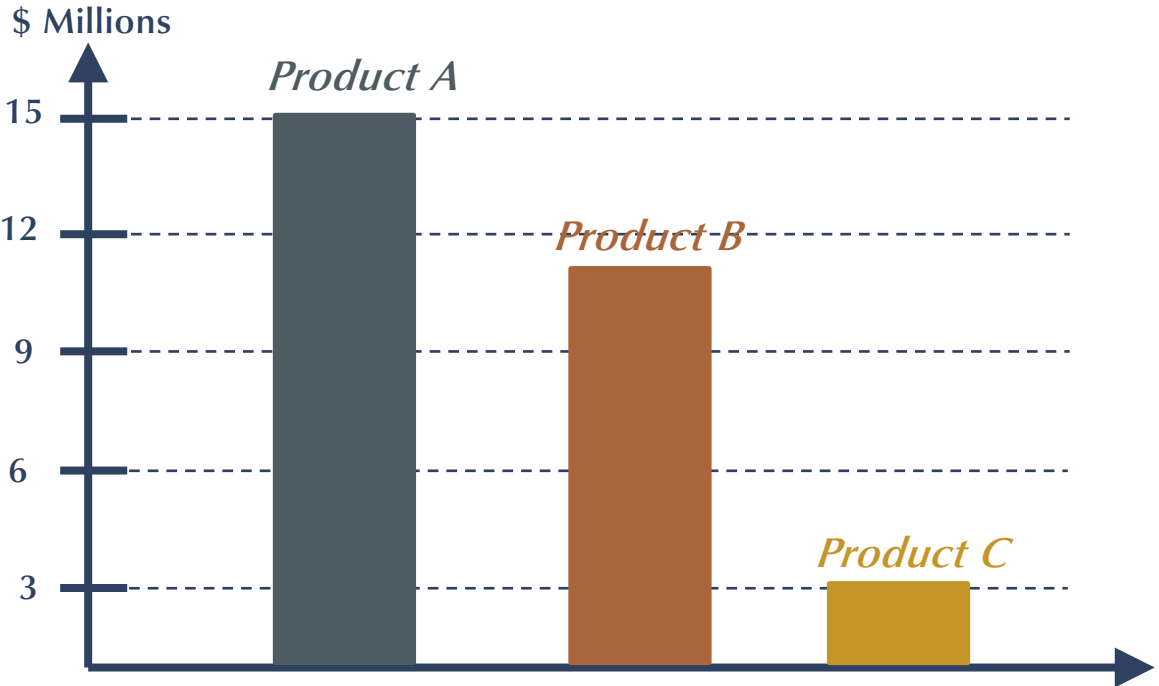
Corrected Charts



Annual Cloud Charges for the Same Data Warehouse Workload with Three Popular Cloud DW Engines

- Real customer database & workload
- Independent consultant
- Different requirement might favor a different product

No one product is always best!!!



Correction: <https://www.beznex.com/wp-content/uploads/2022/02/BEZNext-White-Paper-Which-Platform-is-Best-for-your-Cloud-Data-Warehouse-2-17-2021-1.pdf>

Cloudera Data Warehouse

- Cloudera founded 2008 as a commercial vendor to support Apache Hadoop
- FY21 Company revenue ~\$900 million, growing; lost ~\$200 million
- Acquired by KKR and CDR, now private.
- WinterCorp estimates data warehouse revenues as \$250 million/year, mostly on prem
- Cloudera product line has evolved into the Cloudera Data Platform, now available in public and private cloud as well as on Hadoop clusters on prem
- Data engineering/data science open source focus
- Cloudera Data Warehouse introduced 2017
 - Hive, Impala, Drill are principal query engines
 - Hybrid, multi-cloud emphasis since 2018
 - “Cloudburst” feature enables designated workloads to run in the cloud automatically
 - “Big data” flavor – emphasis on data variety,
 - Supports very large tables
 - Emphasis on high scan rates, analytics, machine learning
- Data architectures are typically highly denormalized

Microsoft Azure Synapse Analytics

- One of the world's most valuable companies: market cap \$1.5 trillion
- Microsoft SQL Server – annual revenues in the billions to tens of billions of dollars
 - One of the three most popular databases worldwide (db-engines.com)
 - Based on Sybase SQL Server (1984) – a leader in the early relational database era focused on transaction processing
 - Deeply integrated with other Microsoft Products
 - Widely used, sound product but historically used more for operational databases;
 - Historically, mainly used for small to midrange data warehousing
- Cloud data warehouse service - Azure Synapse Analytics
 - Cloud data warehouse product introduced 2019: serverless and provisioned options
 - Separates compute and storage
 - Based on surveys of and conversations with of hundreds of cloud data engineers, data scientists and cloud data analysts
 - Goal: analytics at scale in the cloud with good end-to-end integration for development and operation
 - MPP layer over Microsoft SQL Server worker nodes – evolved from Azure SQL Data Warehouse
 - Focused on BI, Analytics, machine learning, includes integration with open source tools (e.g. Spark), includes deep partnership with Data Bricks
 - Supports a broad ecosystem including the data lake and the data warehouse with security model integrated across the whole
 - Similar to SQL Server but not identical
 - Pursuing “HTAP”

Handoff: NK to RW

Oracle Autonomous Data Warehouse (ADW)

- Oracle Corporation Founded 1970's, FY22 revenue \$42.4B, profit \$4.5B, growing
- Oracle Database – Most widely used high end database product – converged database architecture handles relational, JSON, key value, graph, spatial, text, blockchain, etc.
- Autonomous Data Warehouse (ADW) introduced 2018 –
 - Oracle Exadata Hardware now includes persistent memory (PMEM) – much lower access times than SSD
 - Oracle Public Cloud
 - Oracle Cloud@Customer (private cloud on customer premises)
 - Elastic instances – increase/decrease resources with immediate effect – no shutdown
- Exadata Database Cloud Service – “non-autonomous Oracle Database”
- Vanilla Oracle database product available on AWS
- Mature data warehouse product, heavy duty, functionally complete
- “Autonomous” version is much more easily used than traditional Oracle product
 - Virtually no physical database design required, no tuning
 - Still same code base, applications and databases generally migrate as is
- Supports continuous operation; automatic Data Guard now available
- Highly secure, data is completely protected from admins, tech personnel at Oracle
- Hundreds of ADW customers



The End



Richard Winter
richard@wintercorp.com

Norbert Kremer
norbert@wintercorp.com



WinterCorp