ViRTUALiZATiON
& Cloud Review

RE-PLATFORMING DATA WAREHOUSES – WITHOUT COSTLY MIGRATION OF APPLICATIONS

Exploring Adaptive Data Warehouse Virtualization Technology

DATOMETRY

datometry.com

DATABASES NO LONGER AN OBSTACLE TO ADOPTING CLOUD-NATIVE TECHNOLOGY

When enterprises want to move part or all of their data management to the cloud, the biggest problem isn't typically to do with compute or storage, nor is it about performance or latency. The biggest barrier is moving the applications that need to use that data, including industry-specific middleware and business intelligence applications. How does an enterprise move the data management to a new platform without having to re-write all of the applications that rely on the database?

Re-writing entire applications may sound extreme, expensive, and risk-laden, but it has been the traditional solution to database re-platforming initiatives and projects. Over the last few decades, enterprises have built complex disparate suites of applications— point of sales, logistics, analytics, reporting— that communicate with a central database. Unfortunately, these applications can't simply use any database other than the one they were written for originally.

Even if databases use a standardized language such as SQL, the syntax varies greatly from one to another. Functions or stored procedures may be supported in one database and nonexistent in the next. This holds true for every database, including the native database solutions provided by public cloud vendors. What this means for a database re-platforming project is that all the applications must be re-written and tested for the new database before moving to the target database.

An additional consideration confronting enterprise IT is planning and executing a data management migration when the data that needs to be moved is heavily used and frequently updated. It is not unusual for enterprise data warehouses to receive millions of queries from thousands of users in the course of a week. Planning, testing, and executing an enterprise migration to the cloud can be a multi-year project with a high risk of disruption. And, clearly, a multi-year data warehouse migration is the opposite of agility.

The key reason that enterprises are moving into the cloud is the requirement for agile business IT that can take advantage of new technology trends and realize significant savings in CAPEX and OPEX. Typically, provider-specific cloud-native databases are often priced well below what it would cost to light up a VM, install a database inside it, and import the data. And, cloud offerings are evolving rapidly to keep up with changing technology.

So how can enterprises move their data management to the cloud with only a fraction of the cost, time, and risk?

DATOMETRY ADAPTIVE DATA WAREHOUSE VIRTUALIZATION TECHNOLOGY

Technology startup Datometry has a solution: first-of-its-kind Adaptive Database Virtualization technology which enables enterprises to adopt cloud-native databases without having to re-write legacy applications. Despite being an early stage startup, Datometry is getting a lot of attention from global Fortune 500 companies - typically late adopters of technology.

Datometry has also attracted attention from large cloud service providers.

Datometry's flagship product, Datometry® Hyper-Q™, intercepts network traffic between applications and databases and translates it to the language and protocol of the new data warehouse. Datometry is separating the problem of moving data from the problem of rewriting applications for the new database. Doing the translation as a network intercept means there is no need to rewrite and test every application that talks to a database.

DATOMETRY'S FLAGSHIP PRODUCT, **DATOMETRY® HYPER-Q™, INTERCEPTS NETWORK TRAFFIC BETWEEN APPLICATIONS AND DATABASES** AND TRANSLATES IT TO THE LANGUAGE AND PROTOCOL OF THE NEW DATA WAREHOUSE.

Hyper-Q translates the application query text and doesn't do any processing of application queries, which means that the heavy computation is done by the database. Typically, the translation of application queries takes from 5 ms – 200 ms. The database queries are usually running for seconds or even minutes at a time, so the performance overhead is negligible, often undetectable by monitoring.

In addition, if Hyper-Q encounters functions or stored procedures that are present in the existing database but missing in the new database, it emulates them. Stored procedures consist not only of SQL, but of control flow code such as loops. The end result of a stored procedure is typically a series of individual statements that are executed against a database. Hyper-Q unrolls the stored procedures and translates them into individual statements in a syntax that is understood by the receiving database. This allows existing stored procedures to be used even though the target database does not support them.

Hyper-Q currently supports translation from Teradata to Amazon Redshift, Azure SQL DW, Snowflake, , Pivotal Greenplum, and others, with support for other databases planned for Q3 of 2017.

THE STATELESS ADVANTAGE IN VIRTUALIZING DATA WAREHOUSES

Individual Hyper-Q instances are stateless and scalable. This allows Hyper-Q to be highly available and work with existing solutions, even as new instances are spun up to meet growing demand. In this sense, Hyper-Q is similar to other scalable infrastructure components of a public cloud, such as load balancers.

Once the initial characterization and porting has taken place, Hyper-Q instances are plug and play. New instances of applications can be spawned along with Hyper-Q and do not have to be configured to work with additional instances of the target database.

DATA WAREHOUSE MIGRATION WITH DATOMETRY

Migrating a data warehouse requires three key components to be migrated:

- Database schema
- Data, as in the content of the database, and
- Database functions, such as stored procedures

The first step in a Hyper-Q-powered re-platforming effort is getting insights into which application workloads can be re-platformed to the new data warehouse. This analysis is done using Datometry's powerful and automated query analysis software-- Datometry® Hyper-Q™ QueryIntelligence™ (QI) Edition—which generates a full inventory of workload features by analyzing query logs. Using this analysis—which can be completed in hours—data warehouse migration and implementation plans can be created in weeks instead of months.

The reports include information on out-of-the-box coverage of applications, query insights, recommendations for performance tuning and optimization, and a list of database objects referenced by the workload. As a by-product, should any non-translatable constructs be detected, this input will be used by Datometry to guide their development to close any feature gaps.

After the initial characterization is complete the second phase of the migration moves directly to creating cloud instances of the application and testing against the target database.

REDUCED TIME, COST, AND RISK WITH VIRTUALIZING THE DATA WAREHOUSE

Datometry has some real-world customer success stories. The company's first POC involves a Global Fortune 100 retailer looking to move their very large, custom business intelligence application with close to 40 million application queries executed per week to Microsoft Azure SQL DW. Their own testing and POCs had found that to

THE FIRST STEP IN A HYPER-Q-POWERED RE-PLATFORMING EFFORT IS **GETTING INSIGHTS** INTO WHICH APPLICATION WORKLOADS **CAN BE RE-PLATFORMED** TO THE NEW DATA WAREHOUSE.

rewrite the approximately 40 million queries for the new cloud data warehouse would be a multi-year project with costs running in the tens-of-millions of dollars. Datometry's POC was able to demonstrate that Datometry could enable the migration to the new data warehouse within 12 weeks.

THE FUTURE OF DATOMETRY

Datometry has started with data warehousing and cloud migration use cases because it is the strongest market force right now, but Hyper-Q should be useable on private and hybrid clouds as well. The strength of Datometry's Data Warehouse Virtualization technology lies in the fact that it doesn't matter what the database is or where it is located: as long as there is connectivity and Hyper-Q knows the language being used, it will translate query statements and results in real-time.

Datometry will continue adding database support in accordance with customer demand and demonstrated market interest. Datometry believes enterprises should be free to use their applications on the databases that suit them best, and will continue to work to deliver that capability.