modak

As determined by Gartner, through 2022, more than 50% of data migration initiatives will exceed their budget and timeline, impacting business innovation and digital transformation initiatives due to the strategic tools selected and resource skills.

The purpose of this white paper is to focus on the key questions that need to be asked when planning for an enterprise data migration from traditional technology platforms to 'Big Data' cloud platforms, as implemented by AWS, Microsoft Azure, and Google Cloud.

Key Questions

- **Why should a cloud dimension be added to a data management portfolio?**
- **What data should be moved to the cloud?**
- **When should the data be moved?**
- **Where is the data moved to?**
- **How is the data moved?**

Here are some assumptions that provide grounds to discuss these questions.

- Firstly, there is currently an on premise big data platform that needs to be moved to an off-premises hosted cloud provider to reduce cost, improve agility and consolidate data sources.
- Secondly, the enterprise has critical systems and applications that will remain in a traditional environment and may remain on the system as a record for operational and transactional data.
- Thirdly, cost projections and comparisons among various cloud providers have been excluded in this discussion.

modak

Why should a cloud dimension be added to a data management portfolio?

The current on premise hosting of big data platforms presents several challenges which provide the different motivations to adopt cloud technologies for data management.

- Data is spread across different systems that need to be analyzed in a single context and it is not possible to integrate the existing or available tools on the internal hosting infrastructure currently in place.
- The scale of the data has grown too large to be queried and/or analyzed fast enough to support decision making.
- Competitors are adopting cloud technologies at a faster pace and are gaining competitive advantage, or an opportunity has arisen to move faster than your competitors, but unable to capitalize on the opportunity due to lack of tools and skills. Another criterion is the availability of new tools to provide high performance analytics, such as Microsoft Power BI

What data should be moved to the cloud?

Cloud is not there to process all your data from current systems – The cloud has a cost and utilization factor that must be managed very carefully. It does not mean there are no circumstances driven by business goals, cost optimization and agility that warrant a migration to the cloud but these migrations and the ongoing operations needs to align to business objectives.

- Current or Legacy: Enterprise data systems evolve to meet the businesses' changing needs. Often systems and the data generated is historical and is archived and must be retained for regulatory or audit purposes.
- Fit-to-purpose: Not all enterprise data and systems will benefit from migration to the cloud. Some data is more suited for cloud deployment than others.
- Access and Compliance: Compliance with regulations is a critical decision factor in determining what to move to the cloud. Numerous rules and regulations are surrounding of the use of data.

modak

When should the data be moved?

The answer is not ‘immediate’. Technically it is possible to move data very quickly to the cloud. Strategies like ‘Build the data lake and they will swim’ or ‘Just lift and shift’ are strategies that can be implemented quickly with minimal planning.

- A satisfactory set of reasons have been developed when considering the ‘Why’ question. The question is now ‘When’ rather than ‘If’. In other words, there is a strong business case for the migration of data to the cloud.
- There are at least three initial candidates of enterprise data that have been identified for cloud migration. It is accepted that if an enterprise has less than three candidates, then is it not a strong enough business case for migration to the cloud.

Where is the data moved to?

There are numerous cloud platforms available for selection (AWS, Google Cloud, Azure) or even a multi-cloud approach. The ‘Where’ in this discussion is more about the various landing areas in which data will end up.

- Big Data Warehouse.
- The Data Lake.
- The Virtual Cloud.
- The Cold Storage Warehouse.

How is the data moved?

There is no single answer to this question. The analyses and decisions that are made when considering the ‘Why’, ’What’, and ‘Where’ dimensions will determine the technologies that are needed to employ when moving data. However, here are some practices that may be worth considering when migrating a data platform to cloud.

modak

- Re-architecting data platforms
- Separating compute and storage
- Evaluating multiple computing options
- Separating compute and fast data access interface
- Capturing data changes effectively

Modak Nabu™ architecture