LTIMindtree

PoV

Cloud Data Migration Strategy Decoded:

Key Processes, Milestones, and Best Practices

The previous decade saw enterprises storing their data on a few servers in their data centers. Smaller enterprises followed suit and leveraged colocation facilities to run and maintain their servers. But those days are long gone now. Today, enterprises seek modern computing infrastructures that can provide scalability, agility, and data-led insights, which are essential to thriving in a rapidly digitizing business landscape. In fact, Gartner predicts that by 2026, 75% of organizations will adopt a digital transformation model based on the cloud[1].

That's why data migration is more important than ever. It is your organization's ticket to transforming into an agile, scalable, and innovation-ready digital powerhouse. However, getting your data from on-prem systems to these cloud-based data platforms and services is a daunting task, especially when you need to migrate all your data. It is, therefore, not surprising that 83% of data migration projects fail, exceed budgets, or get delayed beyond their schedules[2].

To ensure that your organization becomes a part of the other 17%, it is crucial to guide your migration with a well-formulated strategy. Such a strategy should be comprehensive, adaptive, and watertight to ensure that your digital enterprise is ready to operate on an updated data architecture.

In the coming sections, we'll discuss the different types of migration strategies, factors to consider when devising a migration strategy, and the key steps involved in a migration plan. We'll then deep dive into a software distributor's journey from 10+ disparate data systems to Google Cloud.

Factors to consider when formulating a migration strategy

Here are the key factors that must be considered when formulating a migration strategy:

Data quantity

Knowing the volume of data to be migrated and what techniques will be leveraged to migrate large datasets (parallel processing or compression, for instance).

Data quality

Discovering and addressing data quality, inconsistency, and incompatibility issues that may affect the migration.

Workload type

Databases underpinning key services that require improved performance may receive rearchitecting treatment or upgrade to a newer version.

Apps served

Understand if the applications served by the database can tolerate downtime. Decide the migration approach accordingly and keep the respective stakeholders informed about what to expect.

Target environment

This determines how your upstream processes (like Extract, Load, Transform (ETL)) and downstream processes (like analytics applications) will be affected. For instance, new data warehousing services like BigQuery may be better suited for Online Analytical Processing (OLAP) behavior, whereas services like Cloud SQL will work better for datasets that require Online Transactional Processing.

[1] *Gartner Forecasts Worldwide Public Cloud End-User Spending to Reach Nearly $600 Billion in 2023*

[2] *https://www.oracle.com/a/ocom/docs/middleware/data-integration/data-migration-wp.pdf*

Finally, a data migration strategy should delineate a migration plan that effectively prioritizes the databases to be migrated and the approach that will be employed to actually migrate the data.

An overview of key data migration approaches

Here are the two major migration approaches leveraged to migrate Data Warehouses (DW) to cloud and/or hybrid cloud environments.

Big-bang migration

In this approach, all the data estate of the organization is moved to the target environment all at once instead of in segments.

Pros: Big-bang migrations can be more **cost-effective**. They can **lower operational complexity**, as the old environment is retired right after the migration. Moreover, the actual migration is carried out in a short window and is, therefore, **faster**.
Cons: Big-bang migration is typically considered a **high-risk approach**. Therefore, having rollback controls in place is advisable. Moreover, it may **necessitate a downtime window**.

Trickle migration

In trickle migration, the data is moved to the target system in phases or segments. Multiple datasets may be migrated in parallel, but the overall migration is divided into phases.

Pros: Trickle migration is a **lower-risk strategy**, as failure can be managed effectively. Moreover, databases can be migrated with **zero downtime** for the applications they serve. Finally, trickle migrations offer **better visibility** into the project, especially when massive data estates are involved.
Cons: A trickle migration can be **more expensive** and introduce **complexity** in the form of multiple active data environments. Moreover, the actual **migration process is longer** than a big-bang migration.

Key steps in a data migration plan

Once the migration strategy has been formulated, the migration is initiated as per the outlined plan. Here are the three main steps that are carried out in a migration plan.

Assessment and risk identification

This entails understanding the infrastructure and architecture of the current DW, user and workload requirements, and cataloging the systems within which the data is stored. Risks like data loss, performance degradation, security and compliance issues, and costs of downtime are identified, and a mitigation plan is devised for each of them. Finally, timelines are outlined in the migration plan, along with Definitions of Done (DoD).

Migration of data

This is the stage in which the actual migration is carried out. The data is extracted from the existing DW, then transformed and staged in an intermediary storage environment. Next, the staged data is loaded in the target environment, and database schemas and configurations are migrated to the target system. The migrated data is then tested and optimized for performance. The migration design must specify migration and testing rules, define acceptance criteria, and assign roles and responsibilities across the migration team members.

Validation and governance

Before launching migrated data to production, it is validated through integrity and consistency checks and reconciliation techniques. This ensures that the data has been correctly transported and logged. The results are then validated with key business users. This is the right time to update the metadata repository and review data classifications along with archival and retention policies and controls. After a post-migration audit, the old data warehouse can be retired.

Tips to ensure success with your migration strategy

Some data migration issues surface on the fly and must be tackled by experienced professionals to ensure a successful migration. Over the course of hundreds of data migration projects, we have observed that the following factors are essential to nailing the outcomes of a migration initiative:

01 Brainstorm to **eliminate erroneous assumptions**. For instance, existing data models may not map to the target system, quality issues may be uncovered when migrating the data, and documentation of the data estate may be partial and inaccurate.

02 **Mitigate risk early on** by spending time understanding your data estate. Don't let risks materialize in the form of load failures in your target environment. The longer a risk festers, the more expensive the overall project becomes.

03 **Consider creating a migration factory** that contains repeatable, end-to-end process flows. This creates the necessary guardrails, speeds the project, and de-risks the process by minimizing disruptions to business users.

04 **Invest in change management** early on. Invest in data governance and source technology to enhance data quality, if needed, and start cultivating a data engineering and data management organization before your data is migrated.

05 **Take an iterative approach** if you are addressing multiple issues, such as improving data quality or remodeling your data as a part of the migration project. A big-bang migration is likely to veer off course due to such issues.

06 **Undertake data profiling** before you actually move the data. This will enable you to identify all the systems where the data resides and the quality/integrity issues affecting the data in these systems.

Case in point: Modernizing the data estate of a software products distribution major

We have leveraged the above guiding principles to drive numerous critical data migration projects to success. Through one of our recent engagements, we helped a client modernize their data estate and migrate their data warehouse to Google Cloud – and delivered a significant impact in the process.

The client was in the midst of a data modernization program, with critical data assets housed in 10+ disparate systems. Some were still residing on-prem, whereas others were in flight, in AWS, or other enterprise systems. Our client wanted to unify these data assets and migrate all data workloads to Google Cloud – and consequently engaged LTIMindtree to lead the initiative.

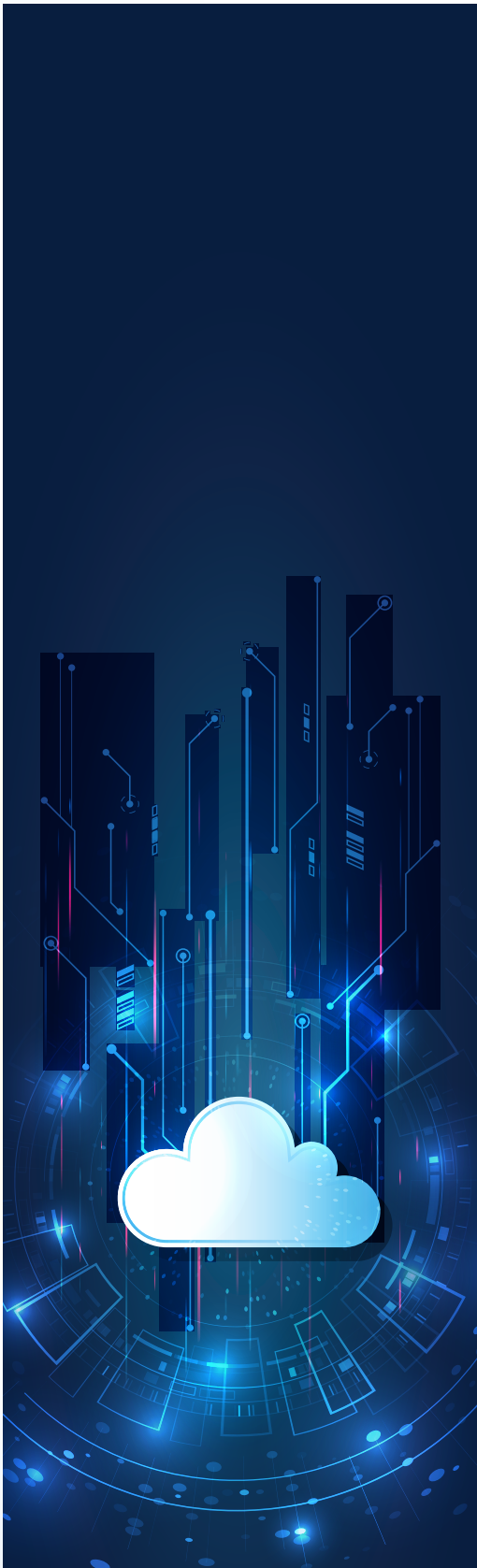
We carried out due diligence on their data estates and identified gaps and failure modes to re-baseline the modernization plan and architectural construct. Following this, we carried out the migration leveraging our proprietary accelerators to deliver 60% automation across the program. The migrated solution was taken live 50% faster, resulting in strategic gains for the client.

Summing it up

Cloud is the inevitable DW destination for most enterprises. The increased scalability, agility, and performance gains enabled by cloud DWs are invaluable in today's digitally driven business ecosystem.

However, any cloud modernization program is bound to fail in the absence of a vetted migration strategy. Since there is no one-size-fits-all approach to migrating DWs, each organization must invest considerable effort into curating a migration strategy that works for them.

The above guiding principles can help businesses carry out successful migrations with repeatable success. However, the assistance of experienced and trusted technology partners can prove invaluable in maximizing the impact of the modernization initiative in the long run.



Authors' profile

Rajat Goswami

rajat.goswami@ltimindtree.com

Rajat has over 12 years of experience in the IT industry across product management and consulting in critical business applications and bleeding-edge technologies, such as data analytics and AI/ML. His unique skillsets and experience allow him to find the right technology solutions for industry challenges. A technology enthusiast at heart and an alumnus of IIM Calcutta, he has conceptualized and developed AI-enabled business products and led complex client engagements from initial stages to delivering value for business. Currently, he is working as a Data and Analytics Consultant at LTIMindtree, driving GTM strategies and solutions for the D&A practice.

Harshavardhan Jb

harshavardhan.jb@ltimindtree.com

Harshavardhan has over 17 years of experience in data and analytics consulting, crafting solutions to enable self-service analytics and data solutions. He is an agile evangelist, proficient in transforming business from the traditional approach to agile ways of working, and has performed various roles both in technology and program management. He specializes in EDW modernizations, enabling clients to move from on-premise data ecosystems to cloud-native systems. He is currently working as a GCP Solution Architect and Technology Consultant for LTIMindtree.

LTIMindtree is a global technology consulting and digital solutions company that enables enterprises across industries to reimagine business models, accelerate innovation, and maximize growth by harnessing digital technologies. As a digital transformation partner to more than 700 clients, LTIMindtree brings extensive domain and technology expertise to help drive superior competitive differentiation, customer experiences, and business outcomes in a converging world. Powered by 83,000+ talented and entrepreneurial professionals across more than 30 countries, LTIMindtree — a Larsen & Toubro Group company — combines the industry-acclaimed strengths of erstwhile Larsen and Toubro Infotech and Mindtree in solving the most complex business challenges and delivering transformation at scale. For more information, please visit **https://www.ltimindtree.com/**