

Modern data management

The key to winning banking customers





Contents

Introduction	3
Moving financial data at transactional speed	4
Addressing AI challenges	7
Securing customer data	7
Leveraging the power of AI	8
Staying ahead of the market	9

Introduction

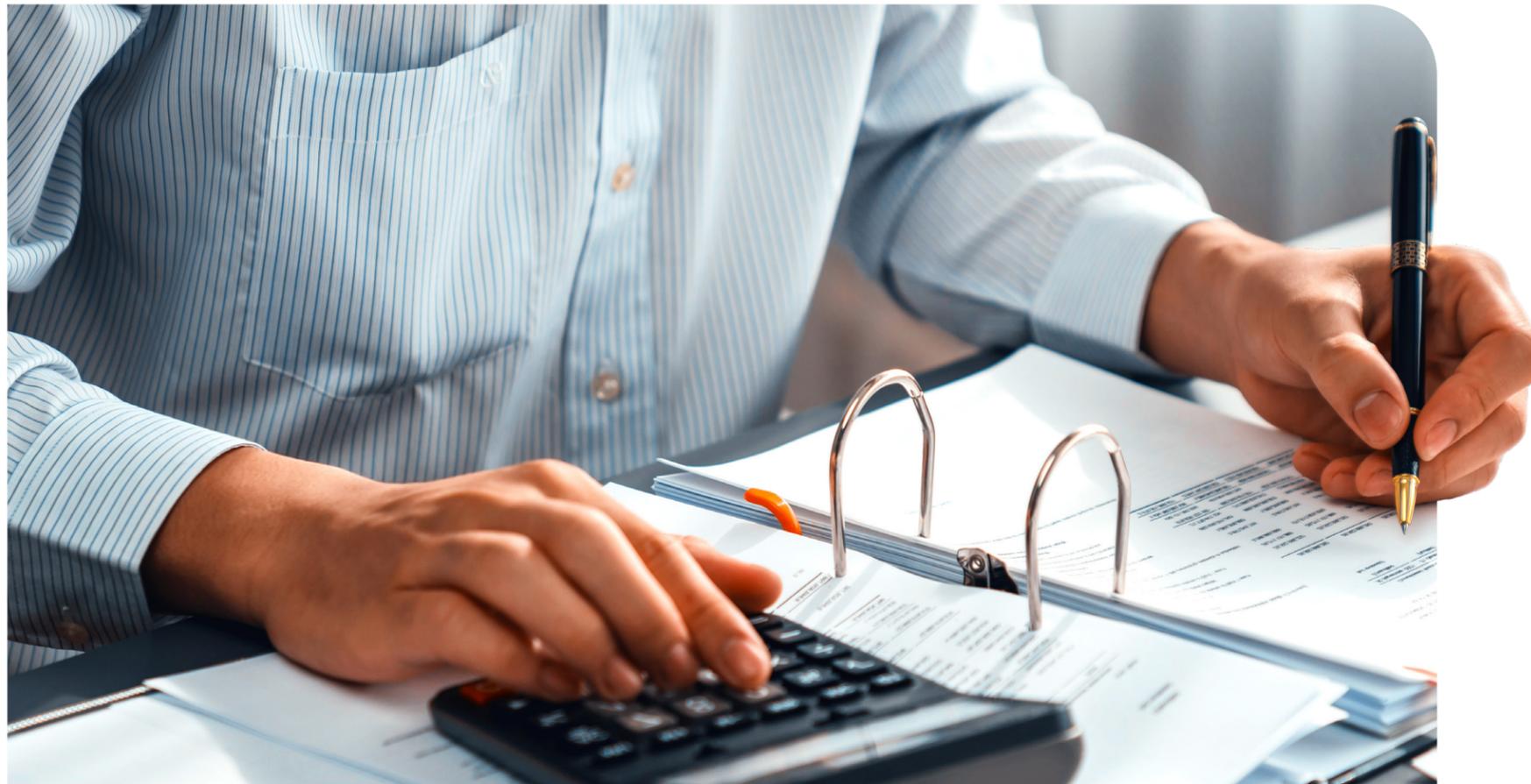
The legacy banking industry is under siege with the emergence of neobanks, digital-first banks that appeal to the next generation of tech-savvy customers who expect all services to be fast and smartphone-friendly. At the heart of this competition for customers is data powered by modern technology. Digital-first financial organizations have no encumbrances from outdated technology infrastructure and, without physical branches, less overhead. It enables them to offer leaner services, notably checking accounts and debit cards, the basics of everyday living. And they can offer these services in the quick, on-demand style their customers want.

How can traditional banks compete more effectively for next-gen customers and retain their existing customers as digital banking becomes the norm? It requires a new approach to data, understanding that modern data management is the key to delivering a customer experience that is even more nimble than the competition's. Updating the data infrastructure that governs data transit and storage will also further ensure banks can meet requirements for data privacy and compliance.

Moving financial data at transactional speed

Each day banks move terabytes of data that is generated by financial transactions, customer inquiries, market data insights, and data analytics. To compete effectively means leveraging this valuable data and responding to customer needs at a pace that matches digital first banks. Banks that are keeping pace are developing new products and services, gaining deeper insights into customers' spending and transaction history to better segment their market, exploring ways to integrate services with other organizations, such as offering cash back incentives, and promoting customer retention by offering customized financial planning and other support services.

Demonstrable value and insights can be gained from the large volumes of data flowing into banks. However, to successfully leverage this data requires a more sophisticated approach to managing the availability of stored data for analysis, model training and other uses.



Banks want to make use of this data, but it cannot come with high costs and unnecessary complexity. A modern data management solution addresses these concerns by providing:

- **Hybrid cloud support:** IT can freely adopt and use the best cloud environment for any fact of operations, benefitting from price and capability differentiation to control costs and improve outcomes.
- **Scalability:** Banks can further control costs by scaling dynamically as needed, maximizing the productivity of available bandwidth.
- **Migration flexibility:** Where data is initially accumulated may not be the ideal environment in which to extract the most value. To be efficient and realize the most benefit from the terabytes of data flowing in daily, banks need data migration technologies that enable the transfer of large-scale data among heterogenous systems, and that support open and interoperable methods for using those datasets. Banks will then be able to move datasets of any size, from on-premises data lakes to the cloud, in less than half the time of typical approaches, without disrupting day-to-day transaction flow.

Technologies like generative AI, language models, and modern data analytics need direct access to data to operate most effectively. Providing data in a usable form automatically makes their use more efficient and accurate, improving the outcomes that can be achieved by adopting AI capabilities.
- **Automated efficiency:** Migration occurs without any need for custom coding. Automation and faster time to value frees up IT time for strategic tasks.

As traditional and neobank customers make digital their preferred way to bank, delivering transactions and information at speed and scale is a key factor in customer satisfaction. A solution that delivers hybrid cloud capability, economical scale options and efficient transfer of datasets between systems provides a solid foundation for competing with leaner neobanks.



Addressing AI challenges

Like other industry sectors, banking institutions are looking at how to integrate generative AI (GenAI) into their data infrastructure. GenAI can add petabytes of data to a banking operation, but it comes with challenges like determining which data is valuable for customer insights and the cost of resultant cloud storage. Also as PwC notes, "As a highly regulated industry in possession of sensitive data, banks must add GenAI model governance to their current AI governance practices, employ data ring fencing and put compliance and reporting at the forefront."

Securing customer data

GenAI adds more data and complexity for traditional or digital-first banks charged with the responsibility of protecting customer privileged information. It opens the door to inadvertent sharing of privacy-governed information. The solution is rapid transfer of large datasets to the appropriate cloud or on-premises location, according to its intended use. More precise control over the data, through organization and secure transfer, must be a paramount objective to avoid a privacy breach.

Secure transfer is just one part of the equation. Continuous replication must be part of any modern migration, ensuring that disaster recovery (DR) can occur efficiently without business disruption. Banking IT teams also must revisit their business continuity protocols to adjust for GenAI data. To achieve near-zero recovery time objectives (RTO) and recovery point objectives (RPO), a solution must provide continuous cloud replication of actively used GenAI data that informs critical applications and projects.

Replication also ensures current data accuracy should recovery be necessary. The development and analytics teams will then be able to continue working with the most up-to-date material.

Leveraging the power of AI

While GenAI can produce enormous amounts of data that offer competitive customer insights, to be valuable it must be managed and made available to other systems.

To achieve this, banks will benefit from a fully automated data migration solution that moves on-premises HDFS data, Hive metadata, local filesystem, or cloud data sources to any cloud or on-premises environment, even while those datasets are under active change.

In addition to customer insights, banks are recognizing how GenAI can improve internal operations, taking the terabytes of information they collect to discover opportunities that enhance work in wealth management, investments, trading compliance and risk management. All these tasks involve confidential data and require data migration that is continuous and secure.

To begin leveraging GenAI, banks are using AI-powered bots to deepen customer engagement. Wells Fargo, for example, has taken bots to the next level with its virtual assistant, Fargo. Part of its mobile app, Fargo understands natural language and wants to be more than just the typical automated bot. Wells Fargo says, "It's more than just dollars and cents; it's about uplifting our customers' emotional and financial well-being by understanding their financial goals and providing the most convenient interactions to meet those goals."

The Fargo bot is a prime example of the need for a data migration solution that can support the large volumes of customer data flowing into the system. It needs continuous data transfer to enable retrieval augmented generation (RAG) to respond to a user query, using GenAI to interact with existing large language models, and arrive at a satisfying answer for the customer at a desirable speed.

Deutsche Bank has gone all-in on GenAI, launching a bank wide, business-driven AI initiative in 2023. They say, "we are leveraging our partnerships with top-tier technology firms, like Google and NVIDIA, to develop AI capabilities spanning content management, anomaly detection, workflow optimization and other use cases to bring new services and capabilities to our clients."

Working in partnership with Google Cloud, Deutsche Bank has built a modern data analytics and AI/ML platform to develop use cases including AI-chatbots, AI tools to generate code for software developers, testing and documentation, as well as the analysis and processing of unstructured data. The bank already has experience using AI to field client queries, analyzing messages coming into group mailboxes, routing the messages, and structuring a request to assure a timely response to the client.

These banks, and others, are seeing the competitive value of GenAI. Coupled with an effective, modern data management architecture they will be able to realize the benefits of additional consumer intelligence and improved internal operations.

Staying ahead of the market

Digital-native consumers, are immersed in mobile technology and expect their banking institutions to be digitally sophisticated. Many will likely never set foot in a physical bank. To capture these customers, banks need to anticipate their needs with new services, getting the customer intelligence to do so from the terabytes of data they collect.

One survey of banking customers noted the top wants: security and fraud protection features, quality customer service, mobile and online access and low checking and other account fees. Leaner, nimble digital first banks are already delivering the type of services customers want. Traditional banks need to become as nimble and harness the power of data, using an efficient modern data management solution to help turn data intelligence into engaged customers.



Case study #1:

NatWest re-architect their Hadoop data lake for the cloud.

Objectives

To be able to take advantage of Sage Analytics within the AWS cloud, NatWest planned to migrate their current on-premise Hadoop based data, the Central Customer DNA Database to the Amazon cloud.

Challenges

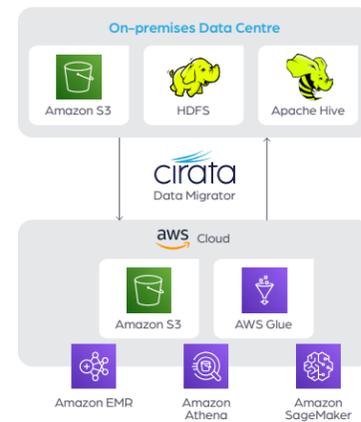
NatWest's on-premises data lake used Hive metadata that they wanted to consolidate into the Amazon Glue repository in the cloud. They also needed to move the results of analysis in Amazon back to on-premises storage to support regulatory reporting applications that had not been adapted for cloud use.

Following a proof-of-concept (PoC) NatWest selected Cirata for their on-premises data lake to AWS cloud data transfer process. Data Migrator is an automated, scalable, high performance, and cloud-agnostic data integration solution that simplifies making data available in and immediately usable across on-premises environments and with any cloud platform. The PoC demonstrated that Data Migrator would meet all of NatWest's requirements and address their data transfer challenges.

NatWest's original solution for moving data to Amazon involved relying on the Cloudera Replication Manager and scripted functions in AWS Lambda. Replication Manager uses the 'DistCp' functionality of Hadoop to move the data, which didn't satisfy NatWest's requirements.

- It is labor-intensive to manually reconcile differences because of data changes made since the last DistCp run. This results in higher costs and often leads to delayed and failed projects.
- Multiple scans are required to capture ongoing data changes made between DistCp runs. Depending on the size of the dataset and the number of changes occurring, it may be impossible to ever catch up with all changes.
- DistCp runs as a standard MapReduce job competing for resources with other processes and requires you to have open firewalls across all nodes in the cluster, posing security issues.

Data Migrator performs the initial data transfer using a single scan of the source storage, while also supporting continuous replication of any ongoing changes from source to target with zero disruption to current production systems.



Results

Data Migrator enabled NatWest to:

- Data flow optimization – Automate and manage the data transfer far easier than their previously script driven solution.
- Provided a standard mechanism for moving data whether it be a cloud target or on-premise.
- Future proofed the solution to take advantage of newer analytics and Metadata storage.
- To date 1.4PB of data moved.
- Increased the number of data science and innovation lab experiments to develop machine learning models across the bank.

Results

Data Migrator enabled Westpac to:

- Automate the data transfer from the on-premises Hadoop cluster to Microsoft Azure.
- Optimize data transfer performance and network bandwidth usage.
- Provide the data to enable upstream applications access to the data in the cloud. This required data movement between ADLS domains.
- 2 PB of data moved to the cloud.

Case study #2:

Westpac begin to exploit their digital and hybrid multi-cloud strategy.

About Westpac

Westpac is Australia's first bank and oldest company, one of four major banking organizations in Australia and one of the largest banks in New Zealand. The bank is halfway through their 5 year program to modernize its technology environment and expand use of cloud-based systems. To achieve this Westpac, have partnerships with both Microsoft Azure and Accenture.

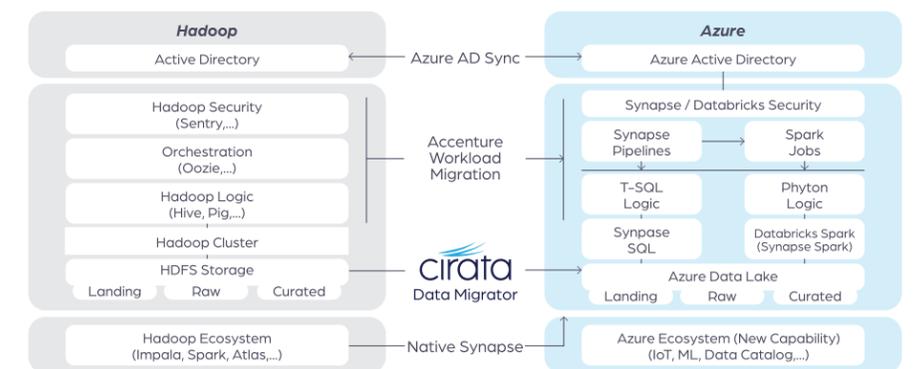
Objectives

To embellish the bank's capabilities in customer insight, the bank wanted to utilize the analytics capabilities of Azure Synapse. To create a quick win for this business case Accenture were commissioned to help provide an ingestion process to move data currently stored on the bank's Hadoop data lake environment into Azure data lake storage. With a significant investment in Microsoft, Westpac needed to show early returns, and their key project drivers were to:

- Provide early access to Microsoft Analytics tools
- Introduce the new capabilities without interrupting existing business processes.

Solution

Cirata Data Migrator was used successfully to move the Hadoop data into the Azure Data Lake, and once migrated also move this data to other domains within the Azure infrastructure. The data migration was a live migration in that any updates that occurred to the Hadoop cluster during data migration were also captured as part of the migration process, ensuring that the cloud version of the data was in sync with the original.



Cirata is your data migration and management partner to win banking customers in the digital marketplace.

Contact us to learn more about Cirata Data Migrator and how we can help you move, integrate, and monetize your consumer data.

About Cirata

Trusted by global brands and industry leaders for more than 15 years, Cirata specializes in the migration of Hadoop data lakes into leading cloud platforms to enable game-changing Artificial Intelligence (“AI”) and analytics. With Cirata, data leaders can leverage the power of AI and analytics across their entire enterprise data estate to freely choose analytics technologies, avoid vendor, platform, or cloud lock-in while making AI and analytics faster, cheaper, and more flexible. Cirata’s portfolio of products and technology solutions make strategic adoption of modern data analytics efficient, automated, and risk-free. In addition, leveraging our patented technologies, including the Distributed Coordination Engine (“DConE®”), our DevOps solutions integrate effortlessly with your existing source code management to increase security, minimize risk, reduce latency, and improve collaboration across globally distributed development teams.

For more information on Cirata, visit www.cirata.com.