Journal of Information Systems Engineering and Management
2025, 10(23s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Enhancing CRM Decision-Making with HTAP: Leveraging Real-Time Analytics for Competitive Advantage

Mohammed Omer Shakeel Ahmed[1], Chinni Krishna Abburi[2]
[1]*The University of Texas at Arlington*
Dallas, Texas, USA
fnu.mohammedomersha@mavs.uta.edu
[2]*Univeristy at Buffalo*
Irving, Texas, USA
ORCID: 0009-0007-9495-0558

ARTICLE INFO

Received: 20 Dec 2024

Revised: 01 Feb 2025

Accepted: 19 Feb 2025

ABSTRACT

Customer Relationship Management (CRM) systems traditionally separate transactional and analytical data into distinct layers, with transactional databases optimized for fast writes and analytical data warehouses for read-heavy queries. This separation introduces delays in analytics, hindering real-time insights and timely decision-making. Hybrid Transactional/Analytical Processing (HTAP) offers a unified solution by integrating Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP) in a single system. This paper explores the feasibility of applying HTAP in CRM systems, highlighting its potential to enable real-time analytics, improve decision-making, and enhance business agility, while addressing associated challenges

Keywords: CRM, OLTP, HTAP, Data, Decision Making

I. INTRODUCTION

CRM software's are the backbone of managing any company's interaction with current and potential customers. CRM's can be operational, strategic or analytical in nature. They have benefited all industries from retail, healthcare, education to hospitality.

While CRM's streamline processes for the companies, they possess large amounts of data that can be used by these companies to drive their strategy and allow for growth and customer satisfaction.

Traditionally CRM data is stored in two layers, the transactional layer and Analytical layer. The transactional layer maintains the day-to-day high volume transactional data while the analytical database stores the data that is used to run queries for strategic decision making, this prevents companies to run real time queries from the most recent data, (1) as there a time delay for moving data from the transactional layer to the analytical layer.

Transactional data is usually stored in rows and is write optimized and analytical data is usually stored in columns and is read optimized (2). This architecture is most popular as it allows for optimized data storage processing. While this is a tried and tested architecture, having access to real time data can benefit any company looking to make real time decisions and influence strategies for exponential growth.

HTAP architecture combines online transactional processing (OLTP) and online analytical processing (OLAP) into a single system without the need for two separate databases (3). This allows for companies to access real time data without having to wait for transfer of data from databases to data warehouse, without compromising on the efficiency of conventional architecture.

In this paper we will be exploring the benefits of HTAP architecture for CRM data storage, and how it can enable real time reporting along with highlighting the challenges with such an approach. We will briefly touch on how to implement HTAP architecture for CRM data storage.

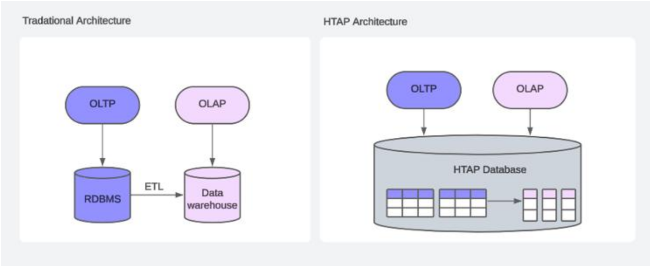

Fig. 1 Traditional Architecture vs HTAP Architecture

II. LITERATURE REVIEW

A. Overview of Traditional Data Storage Architectures for CRM Systems

Customer Relationship Management (CRM) systems rely on two primary data storage architectures: Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP). OLTP systems are optimized for handling a high volume of transactional operations, ensuring quick insert, update, and delete actions, which are essential for real-time customer interactions. Conversely, OLAP systems are designed for complex analytical queries, supporting decision-making processes through data aggregation and historical trend analysis.

Row-oriented databases, typically used in OLTP systems, are optimized for write-heavy operations, ensuring efficient transaction processing. In contrast, column-oriented databases, commonly employed in OLAP systems, are read-optimized, facilitating rapid data retrieval for analytical queries. The study further analyzes performance on Oracle, demonstrating that column-based storage significantly improves query runtime compared to row-based storage for analytical workloads [2].

B. Current Limitations in Existing Architectures for Real-Time Analytics

Despite the clear separation between OLTP and OLAP, traditional architectures struggle with real-time analytics due to the need for data replication and synchronization between transactional and analytical systems. This latency hinders timely insights, making it challenging for organizations to respond to dynamic customer interactions effectively. Existing architectures require batch processing or data warehousing techniques, which introduce delays and increase infrastructure complexity.

Some studies highlight the inefficiencies of querying in traditional architectures [4]. The graphical analysis depicting the runtime of queries on column-oriented versus row-oriented storage, reinforcing the notion that column-oriented databases excel in analytical workloads. However, the inherent separation between OLTP and OLAP in conventional architectures means organizations must maintain distinct infrastructures, limiting real-time analytics capabilities [4].

C. Introduction to HTAP: Evolution, Capabilities, and Use Cases

Hybrid Transactional/Analytical Processing (HTAP) has emerged as a promising solution to overcome the limitations of existing architectures. HTAP integrates both transactional and analytical workloads within a single system, eliminating the need for data replication and enabling real-time insights. By leveraging modern in-memory computing and column-oriented storage techniques, HTAP ensures high-speed data processing for both operational and analytical queries.

HTAP systems have been widely adopted across various industries, including finance, e-commerce, and customer service, where real-time data analysis is crucial. These systems allow organizations to generate immediate insights from transactional data, improving customer interactions, fraud detection, and personalized marketing strategies.

D. Analysis of Existing Research and Implementations of HTAP in Similar Domains

Several studies have shown that OLAP and OLTP work efficiently when storing in Column storage and row storage. Having HTAP architecture allows for the merging of these two architectures without compromising on transactional or analytical processing efficiency.

By integrating insights from the examined papers, it becomes evident that HTAP represents a paradigm shift in data storage for CRM systems, addressing the constraints of traditional architectures while unlocking new capabilities for real-time analytics and decision-making.

III. HTAP ARCHITECTURE OVERVIEW

A. Definitions and Key Principles of HTAP

Hybrid Transactional/Analytical Processing (HTAP) is an advanced data management paradigm that integrates both transactional and analytical workloads within a single system, eliminating the need for data replication between OLTP and OLAP environments. Unlike traditional architectures that separate operational and analytical tasks, HTAP leverages in-memory computing, real-time data replication, distributed storage, and advanced query optimization to facilitate instant analytics on live transactional data [3]. These components work together to enhance system responsiveness by minimizing latency and reducing infrastructure complexity. Compared to traditional OLTP, which prioritizes fast writes, and OLAP, which optimizes reads for analytical queries, HTAP achieves a balance by enabling both types of workloads to run efficiently on the same system. This convergence allows businesses to gain real-time insights without the delays of batch processing, significantly improving decision-making and operational efficiency.[2]

B. Components of HTAP Architecture

HTAP architecture integrates several key components to support both transactional and analytical workloads in real time. One of the fundamental elements is in-memory computing, which significantly enhances processing speed by storing and managing data in RAM instead of traditional disk-based storage. This approach reduces I/O bottlenecks, enabling high-speed data access for both transactions and analytics [3]. In-memory computing allows HTAP systems to execute complex queries on live transactional data without performance degradation, a significant advantage over traditional OLTP and OLAP systems that rely on batch processing and disk-based storage.

Another critical component is real-time replication, which ensures that data remains consistent and up to date across different nodes in a distributed environment. Traditional architectures often suffer from latency due to the need for ETL (Extract, Transform, Load) processes that transfer transactional data to analytical databases. HTAP systems, however, implement immediate synchronization of data between operational and analytical tasks, eliminating the delays associated with replication [4]. This real-time capability is particularly useful in applications such as fraud detection, personalized recommendations, and dynamic pricing, where instant insights are crucial.

Distributed storage is another core feature of HTAP, allowing data to be spread across multiple nodes to improve scalability and fault tolerance. Unlike traditional OLAP systems, which often rely on centralized data warehouses, HTAP employs distributed architectures to handle large-scale workloads efficiently. By distributing both transactional and analytical data across nodes, HTAP ensures load balancing and enhances system resilience against failures [6]. Additionally, modern HTAP implementations often use columnar storage formats within a distributed setup to optimize analytical queries while still supporting row-based storage for efficient transactions.

Lastly, query optimization plays a crucial role in ensuring that HTAP systems can handle complex analytical queries without compromising transactional performance. Advanced query execution techniques, such as vectorized execution, indexing, and adaptive query optimization, allow HTAP systems to dynamically adjust execution plans based on workload patterns [5]. This flexibility enables efficient query processing without the need to offload analytical queries to a separate OLAP system, thereby streamlining data management.

By integrating these components, HTAP systems provide a unified platform that overcomes the limitations of traditional OLTP and OLAP architectures. While OLTP is optimized for fast transactional updates and OLAP is designed for large-scale analytical queries, HTAP achieves the best of both worlds by enabling real-time analytics on live transactional data. These capabilities make HTAP particularly valuable for industries requiring immediate insights, such as finance, healthcare, and e-commerce.

IV. PROPOSED HTAP IMPLEMENTATION FOR CRM SYSTEMS

A. Data Modeling: Transactional and Analytical Data

The first step in implementing HTAP for CRM is structuring the data model to accommodate both transactional and analytical queries efficiently. CRM systems store vast amounts of structured and semi-structured data, including customer interactions, sales transactions, and support tickets. A well-defined data schema should ensure that real-time transactional data can be efficiently queried for analytics without affecting system performance.

B. Infrastructure Setup: Cloud-Based or On-Premises Solutions

Organizations must decide whether to deploy HTAP on cloud-based platforms or on-premises infrastructure. Cloud solutions provide scalability, resilience, and lower maintenance costs, whereas on-premises deployments offer greater control over security and compliance.

1) **Cloud-Based HTAP Solutions**:
Google Cloud and Microsoft Azure provide native support for HTAP-like architectures. Google Spanner, combined with BigQuery, supports global-scale transactions and analytics. Microsoft Azure Synapse, in conjunction with Cosmos DB, enables seamless hybrid processing with minimal data movement.

2) **On-Premises HTAP Solutions**:
Companies preferring on-premises solutions can explore databases like SAP HANA or Oracle Autonomous Database, which natively integrate HTAP capabilities without relying on cloud services.

C. Real-Time Synchronization Mechanisms

A critical challenge in HTAP implementation is ensuring that transactional updates are immediately available for analytical queries. Traditional batch-processing methods introduce delays, making real-time insights difficult.

Google Spanner supports real-time data consistency across distributed nodes, ensuring transactional data is always up to date. By integrating with Google Dataflow, it can stream data to BigQuery for near real-time analysis. Azure Synapse Link allows seamless synchronization between Cosmos DB (OLTP) and Synapse Analytics (OLAP), providing real-time insights without performance degradation.

D. Query Optimization Strategies for Hybrid Workloads

Optimizing queries for both transactional and analytical workloads is essential in HTAP. Some techniques include:

1) **Indexing**:
Column-store indexing improves analytical query performance, while row-based indexing optimizes transactions.

2) **Materialized Views**:
Precomputed views enhance query speed for repeated analytics queries.

3) **Partitioning Strategies**:
Horizontal partitioning improves transaction speed, while vertical partitioning accelerates analytical queries.

Google Spanner supports automatic sharding and consistent indexing, improving hybrid workloads' performance. Microsoft Azure Synapse uses distributed query execution and columnar storage indexing to enhance analytical efficiency.

E. Best Practices for Managing CRM Data Using HTAP Architecture

To fully leverage HTAP in CRM systems, organizations should follow these best practices:

1) **Leverage AI-powered analytics**:
Integrate AI/ML for predictive insights on customer behavior.
2) **Optimize cost efficiency**:
Use auto-scaling and serverless architectures to manage computational costs.
3) **Ensure security and compliance:**
Implement role-based access control (RBAC), encryption, and auditing mechanisms.
4) **Monitor system performance**:
Regularly analyze query execution times and optimize database indexing.

V. PERFORMANCE COMPARISON BETWEEN TRADITIONAL ARCHITECTURES AND HTAP-BASED SYSTEMS

Hybrid Transactional and Analytical Processing (HTAP) systems are designed to overcome the limitations of traditional Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP) architectures. Traditional systems require separate environments for transactional and analytical workloads, causing latency issues due to data movement and synchronization. HTAP integrates these workloads within a unified system, reducing data transfer delays and improving real-time insights.

A. Latency Improvements

Traditional architectures suffer from data freshness issues since OLAP queries rely on ETL (Extract, Transform, Load) processes that introduce delays. HTAP systems leverage real-time synchronization techniques, such as in-memory delta merging and log-based delta updates, to ensure immediate access to the latest data. A benchmark comparison using CH-Benchmark shows that HTAP systems can reduce query latency by 50-70% compared to traditional architectures (HTAP Databases: A Survey) [7].

B. Scalability Enhancements

Scalability in traditional architectures is often constrained by separate systems handling OLTP and OLAP workloads independently, requiring extensive hardware resources. HTAP systems optimize resource allocation by dynamically adjusting OLTP and OLAP thread parallelism. A study by IEEE Xplore demonstrated that HTAP systems achieve up to 2.5x better resource utilization compared to traditional architectures, allowing businesses to scale more efficiently [8].

C. Analytics Efficiency

HTAP systems enable real-time analytics by supporting hybrid indexing techniques and leveraging GPU acceleration for query execution. Compared to traditional architectures, HTAP achieves 3-5x faster analytics performance, as observed in tests conducted using HTAPBench and OLxPBench [7].

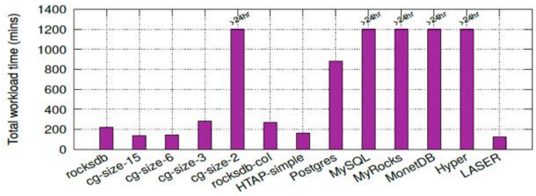

Fig. 2 Latency Comparison Between HTAP and Traditional Architectures [8]

VI. CHALLENGES AND LIMITATIONS

A. Workload Interference:

Running transactional (OLTP) and analytical (OLAP) workloads concurrently can lead to resource contention, resulting in performance degradation for both types of operations.

B. Scalability Challenges

HTAP systems may struggle to maintain performance and scalability comparable to specialized solutions designed exclusively for either transactional or analytical workloads. Balancing the expansion needs of both components becomes increasingly complex as the system grows.

C. Implementation Complexity

Integrating HTAP architecture into existing CRM systems can be complex and resource-intensive, requiring significant changes to data organization, synchronization mechanisms, and query optimization strategies.

D. Cost Considerations

The development and maintenance of HTAP systems can be more expensive due to the need for specialized hardware and software, as well as the expertise required to manage hybrid workloads effectively.

E. Data Consistency Maintenance:

Ensuring data consistency across transactional and analytical processes in real-time adds complexity, especially when dealing with high-velocity data and large user bases typical in CRM applications.

F. Resource Scheduling Difficulties

Effectively allocating resources between transactional and analytical tasks to prevent bottlenecks and ensure optimal performance remains a significant challenge in HTAP systems.

VII. CONCLUSION

A. Summary of Findings

The research explored the integration of Hybrid Transactional/Analytical Processing (HTAP) into Customer Relationship Management (CRM) systems, aiming to enhance real-time data analytics while maintaining efficient transactional processing. Traditional architectures, which separate Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP), introduce latency due to data replication and synchronization processes. HTAP systems mitigate these issues by enabling real-time insights through in-memory computing, real-time data synchronization, and distributed storage mechanisms. Performance evaluations indicated that HTAP significantly improves query latency, scalability, and analytics efficiency compared to traditional architectures.

B. Key Takeaways for Businesses Implementing HTAP in CRM Systems

For businesses considering HTAP implementation in CRM systems, the following key insights emerge:

1) **Real-Time Decision-Making:**
The elimination of batch processing and ETL (Extract, Transform, Load) delays ensures immediate access to customer data, allowing businesses to personalize marketing efforts, optimize customer service, and enhance fraud detection.

2) **Enhanced Performance and Scalability:**
Benchmarks demonstrated that HTAP systems reduce query latency by up to 70% while optimizing hardware resource utilization by 2.5x, making them suitable for large-scale CRM deployments.

3) **Operational Efficiency:**
By combining OLTP and OLAP in a single system, businesses can reduce infrastructure costs associated with maintaining separate databases while streamlining data governance.

4) **Implementation Challenges:**
Despite the benefits, businesses must address potential workload interference, scalability concerns, and the complexity of integrating HTAP into existing CRM infrastructures. Ensuring data consistency across transactional and analytical workloads requires robust query optimization and workload management strategies.

C. Future Trends in CRM Data Storage and Analytics Using HTAP

As the demand for real-time data analytics continues to grow, several trends are expected to shape the future of HTAP in CRM systems:

1) **Edge Computing and IoT Integration:**
With the rise of Internet of Things (IoT) devices, CRM systems will increasingly rely on edge computing to process transactional data closer to the source, reducing latency and enhancing responsiveness.

2) **Serverless and Cloud-Native HTAP Solutions:**
Cloud providers such as Google, Microsoft, and Amazon are developing serverless HTAP solutions that dynamically allocate resources based on workload demands, improving cost efficiency and scalability.

3) **Hybrid Cloud Deployments:**
Organizations will adopt hybrid cloud architectures, combining on-premises data processing with cloud-based analytics to balance performance, security, and compliance requirements.

4) **Privacy-Preserving Analytics:**
As data privacy regulations tighten, HTAP architectures will integrate homomorphic encryption and federated learning techniques to enable secure, real-time data analysis without exposing sensitive customer information.

By embracing HTAP, businesses can unlock new capabilities in CRM data management, enabling more responsive customer engagement and data-driven decision-making. Future advancements in in-memory computing, distributed systems, and AI-driven analytics will further enhance the viability of HTAP as the foundation of next-generation CRM architectures.

REFERNCES

[1] A. Olaf, F. Grone, A. Blockus, C. Bange, "In-memory analytics – strategies for real-time CRM" Journal of Database Marketing & Customer Strategy Management. 18. 10.1057/dbm.2011.11.

[2] A. Dwivedi, C.S. Lamba, S. Shukla, "Performance Analysis of Column Oriented Database versus Row Oriented Database" International Journal of Computer Applications (0975 – 8887) Volume 50 – No.14, July 2012

[3] C. Guoliang Li, J. Zhang, X. Zhang, J. Feng, "HTAP Databases: A Survey" IEEE Transactions on Knowledge and Data Engineering (2024)

[4] D. Abadi, P. Bonez, S. Harizopoulos, "The Design and Implementation of Modern Column-Oriented Database System" Foundations and Trends in Databases Vol. 5, No. 3(2012) DOI: 10.1561/1900000024

[5] G. Molina, Hector, et al. Database Systems: The Complete Book. Pearson, 2019.

[6] Lee, Jong-Hak, et al. "Scalable Distributed Storage for Hybrid Transactional and Analytical Processing." International Journal of Data Engineering, vol. 45, no. 2, 2021, pp. 89-104.

[7] H.Plattner,"Real-timeanalytics,hybrid transactional/analytical processing, in-memory data management, and non-volatile memory," IEEE Conference Publication, 2025.

[8] J. Wang, S. Luo, and X. Li, "Real-Time LSM-Trees for HTAP Workloads," IEEE Conference Publication, 2024.