The Need for Real-Time Analytics

TiDB provides the ability to analyze data in real-time, meaning that you are making operational decisions with the most current and accurate data. This enables you to process and query new data as it is created to inform decisions and guide your business decision-making.

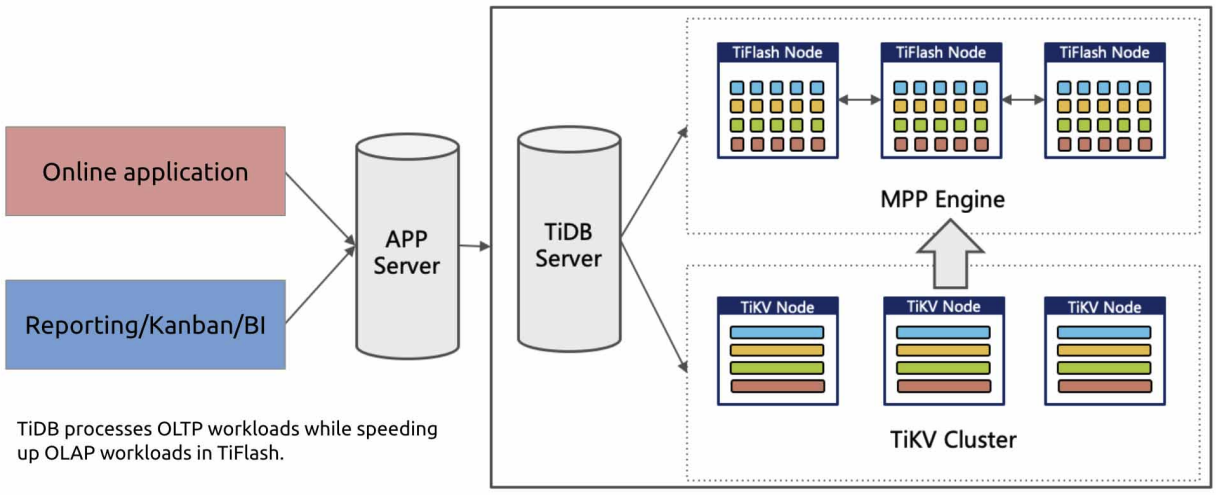
Key Benefits

With the ability to explore your data in real-time, you can adapt your business and its response quickly and appropriately. This leads to

- Enhanced usage of resources in your database environment
- Better business decisions
- Improved customer experience
- Reduction in workload and errors that may come from data transfer

Example Architecture

TiDB handles both Online Transaction Processing (OLTP) and Online Analytical Processing (OLAP) requests. From a user perspective, all data goes to one place and all queries are submitted in the same way. The App Server receives all requests and sends them to the TiDB Server, which dispatches the requests to different storage engines. The architecture is simple, elegant, and easy to implement and use.



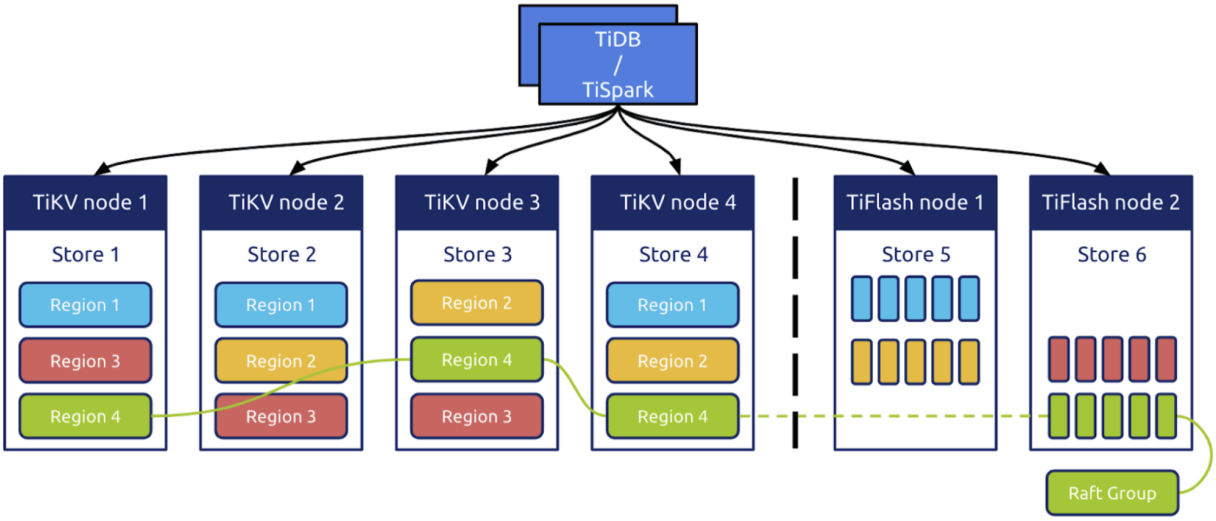
TiDB initially processes a write request to the row store. From there, it is replicated to the column store. Data is written first to TiKV, which is the row store. TiKV provides the speed needed for massive OLTP. Next, the data is replicated to TiFlash, the column store. Data replication is peer to peer with no in-between layer, so the data is replicated in real-time.

Within TiKV and TiFlash, the data is written to a storage location called a region. A TiDB region is not the same as a region within a cloud provider's environment; within TiDB, a region is a storage location. Each region has multiple replicas spread across the nodes of the TiDB environment. Through this replication, TiDB ensures the availability of the data, even when an entire storage location is unavailable. By default, each TiDB region has 3 replicas. High availability is maintained through the implementation of the Raft algorithm, which is a consensus-based method for ensuring the continued availability of data. With a minimum of 3 replicas of each region, spread across nodes of the environment, TiDB can sustain the loss of an entire node without service or data interruption. Both TiKV and TiFlash use the Raft algorithm to maintain data availability.

So, the same data is available for querying in both the row and the column store. The query optimizer determines whether to run each query against the row or column store. It treats the column store as a special index. Among all the indexes in the row store and the special column store index, the optimizer selects the fastest index through statistics and cost-based optimization. This way, both the column and row stores are taken into consideration and the optimizer makes the best decision for each query.

The architecture diagram shows 4 TiKV nodes and 2 TiFlash nodes. Within the 4 TiKV nodes, data availability is ensured since each region replicates data to 2 other regions. This means that, even if an entire node fails, the data is still available for writes and reads. TiFlash does not require multiple replicas since each column store can be recreated from the associated row data, however, it is possible to implement this manually.

For example, the data in Region 4, identified in green, is present on Stores 1, 3, and 4 as part of the row store, and on Store 6 as part of the column store. If TiKV node 3 fails, the row store data is still available via Store 1 and 4 for row-based data and Store 6 for column-based data. The data on the TiKV node 3 will be rebuilt when the node rejoins the cluster.



TiFlash uses a Massively Parallel Processing (MPP) engine. MPP is a storage structure designed to handle multiple operations simultaneously by several processing units. Each processing unit works independently with its own operating system and dedicated memory, meaning each node in a cluster works independently, consumes resources (CPU, Memory, Storage) from its own nodes, processes data on its own nodes, and does not share any data or resources with other nodes.

When dealing with OLAP workloads, TiDB steps back to be a master node. The user sends a request to the TiDB Server, and all TiDB servers perform table joins and submit the result to the optimizer for decision making. The optimizer assesses all the possible execution plans (row-based, column-based, indexes, single-server engine, and MPP engine) and chooses the optimal one.

Example Applications

Businesses need to move at the speed of data, but many organizations are making critical decisions based on data that can be days, weeks, or even months old. History is important, but you must spot trends early and react quickly. Therefore, it is imperative that the data being analyzed is current and reflects activities in real-time. As an example, using analytics to reduce customer defections by 5% can increase profitability from 25% to a whopping 125%.

In a traditional database world, a company processes incoming transactions through a standard OLTP database. This could be MySQL, PostgreSQL, Microsoft SQL Server, or any of a number of other databases. Some organizations copy the data in the database of record to another database, like Amazon Redshift or Snowflake, for analysis using an Extract, Transform, Load (ETL) process. Other companies use the current data in the existing database to do some

analytical queries, but these databases are designed for transaction processing, not analytics, so the queries run slowly and the database resources are not well used.

While both options provide you with analytical capabilities, it would be better if there was a way that data could be immediately made available, in both transactional and analytical structures, at the time that it is initially written to the database. Then, a query is guaranteed to be running against the most current data available, no matter whether it is running as a transactional query or as an analytical query, and it uses the environment's resources efficiently.

Consider a gaming company that has just released a new game. In order to win their users over, they need to be able to make up-to-the-minute adjustments to the game and the leaderboard. The company must ensure that all activity is recorded accurately and promptly so that gameplay can adapt to the user's interactions. The leaderboard needs to be kept current and correctly reflect activity as it is taking place. Without these updates, the company risks alienating and potentially losing customers.

In financial services, delays of even a second can be costly. The cost difference between buying or selling a commodity can vary within milliseconds, so it is crucial that real-time analysis is being done to ensure that transactions are processed correctly. Once a transaction is requested, it is also vital that it be completed in a timely manner to ensure accuracy. A customer's record of activity must show the most recent transactions, otherwise, they lose confidence in your organization.

Another application in the world of finance is fraud detection. Each time a credit card transaction is processed, the system assesses whether or not that transaction matches a user's expected behavior. For example, if the same credit card is used to make a purchase at a physical store in Massachusetts and then just 15 minutes later, the same card is used at a physical store in California, the second transaction is flagged as potentially fraudulent. Even if that second user is the correct user, they are now made aware of the earlier transaction and can close the account to prevent additional fraud.

The Internet of Things has led to an explosion of data. Everything from your car to your refrigerator may be storing and sending content to a manufacturer. The manufacturer can use that data to refine use cases, predict failure, and deliver an improved experience to their customer. However, finding out that your car's battery is likely to fail 10 minutes after it has failed does not help the consumer or the manufacturer. Up to the minute analysis of data, with subsequent reporting and action, is critical to the success of many of today's businesses.

In all of these cases, the ability to process and query data in real-time leads to an improved outcome for the user. There are also benefits to each company that go well beyond keeping their customers happy. On-demand scalability ensures that the company is using its computing resources in the most efficient manner possible. As more organizations move to the cloud, this flexibility becomes more available, but it often comes with a high cost. If a company only scales

up, they may end up paying for resources that they only need during periods of high demand. Some companies, like the gaming company discussed above, can predict times of high usage, since they anticipate a lot of users during and shortly after the launch of a new game. But, for the financial services company, market conditions often drive activity, so they are less able to accurately forecast needs. Scalability on demand is crucial to meet the needs of the consumer and the business provider.

Errors are reduced by eliminating the need for an ETL process to move data. When such an activity is needed, it is possible that it is delayed due to downtime issues, user error, or more. By eliminating this intermediate step, businesses are ensured that the data is current, accurate, and immediately available. This also leads to better usage of personnel resources since the ETL process no longer needs to be run, monitored, and managed, freeing up your people to work on other tasks.

TiDB provides a single source for your data and enables you to query, scale, and adapt in real-time. This means that you keep both your customers and your coworkers happy. Data updates process quickly and query response is speedy. Add the elimination of tedious ETL processes, and your team is better able to work on other, more productive tasks.

Customer Examples

Bigo

Bigo is a social networking company based in Singapore. They have more than 400 million active users each month in over 150 countries. Bigo was experiencing difficulties with scalability and had poor performance with complex analytical queries. They were able to service all of the transactions that were generated by their users but were relying on a batch process to move the data to an analytics platform. With TiDB and TiFlash, Bigo is now able to execute real-time business requests, with execution time reduced to seconds instead of minutes or even hours. There are no more data exports and imports to and from their big data platform. Bigo relies on the TiDB HTAP solution for both their OLTP and OLAP needs. This enables users to anticipate continuous access, even in the face of unexpected usage surges.

Zalopay

Zalopay, based in Vietnam, is a mobile payment application with over 100 million active users. The application enables users to transfer and collect payments with an online chat platform, book hotels and travel, order products, share bills in group chats, and pay through a merchant’s official account. The merchants cover multiple industries such as retail, catering, service, and e-commerce. Their business decision-makers wanted to get insights into their various issues in real-time and understand trends, such as risk management and fraud detection. They also

wanted to expand the use of TiDB analytical services to conduct large amounts of data mining for quick business recommendations. Combining TiFlash with TiSpark, an Apache Spark plugin that works with the TiDB platform to answer complex OLAP queries provides Zalopay the performance and scalability they need. Users get the performance they expect and merchants are happy to receive funds quickly and accurately while reducing fraud.

ZTO

ZTO is the largest logistics company in China. They moved their data out of a legacy big data application to TiDB and now are able to store three times more data at a lower cost, plus they gained a 5x performance improvement in query response. Their live tracking parcel module streams data to TiDB, and the mobile application and real-time reporting is able to show tracking status in real-time.

Additional Resources

Easily Build Your Mission-Critical Applications at Any Scale
A One-Stop HTAP Database Solution
Empower Your Business with Big Data + Real-time Analytics in TiDB