

Distributed SQL

An Enabler for Globally Consistent Transactional Workloads

COMMISSIONED BY



yugabyteDB

MARCH 2020

©COPYRIGHT 2020 451 RESEARCH. ALL RIGHTS RESERVED.

About this paper

A Pathfinder paper navigates decision-makers through the issues surrounding a specific technology or business case, explores the business value of adoption, and recommends the range of considerations and concrete next steps in the decision-making process.

ABOUT THE AUTHOR



JAMES CURTIS

SENIOR ANALYST, DATA, AI & ANALYTICS

James Curtis is a Senior Analyst for the Data, AI & Analytics Channel at 451 Research. He has had experience covering the BI reporting and analytics sector and currently covers Hadoop, NoSQL and related analytic and operational database technologies.



Executive Summary

There is a good deal of talk about the digital transformation journey that enterprises often embark on as a means to evolve as a modern software-defined business. 451 Research finds that an overwhelming majority of enterprises are either actively pursuing digital transformation initiatives or are in the planning phase. The rest are trying to play catch-up and have little, if any, plans for digital transformation initiatives.

Because all enterprises use and manage some type of database system—often transactional systems— digital transformation initiatives are likely to impact these systems, as well as other areas of the organization. Relational database systems that drive transactional workloads are expected to remain popular. But the idea of disrupting these systems gives many organizations great angst, because transactional systems often hold the company's most critical data, and any impact to that system will directly impact the business. That leaves enterprises realizing that transactional workloads will continue, even though the underlying technology may change. What are enterprises to do?

Relational databases generally drive transactional workloads, primarily because of their ability to leverage SQL's data modeling/query flexibility, as well as ensure absolute data integrity through guaranteed ACID transactions. Although other types of data platform systems can handle transactional workloads, research suggests that the SQL query language will remain the lingua franca of relational systems. That means enterprises can find solace in the fact that despite a potential technology change at the database level, core SQL skills will remain as a means to manage and manipulate the data within relational systems.

One of the expected changes for transactional systems will be where they run. Historically, transactional systems have run on-premises, but research points to cloud as the primary location of transactional workloads going forward. But getting to the cloud can be a complex journey, especially when organizations are looking to transition legacy applications and systems. The fact is, simply 'lifting and shifting' to the cloud could have long-term ramifications because the data platforms and applications have not been optimized to fully leverage cloud infrastructure, and would thus not run in an optimized state.

Ideally, enterprises need to take the long view and seek to match their data platform systems with cloud-native infrastructure. To that end, there are relational database technologies that have been specifically developed and architected to run natively in the cloud, taking advantage of the inherent characteristics of all that the cloud offers. Such characteristics include heterogeneous commodity infrastructure, near-instant elasticity, the promise of unlimited capacity and the ability to deploy across multiple datacenters around the globe. Research suggests that when enterprises can find maximum value from their investments—digital transformative initiatives, for instance—they can benefit and run their organizations with optimal operational efficiency.



Database Systems Sit at the Heart of the Enterprise

Digital transformation has become somewhat of an industry catchphrase. But the underlying intent behind the concept is to address certain gaps within the organization. 451 Research has found that digital transformation efforts are alive and well, with 70% of enterprises either actively executing (43%) on digital transformation projects or in the throes of evaluating, planning and researching their digital transformation strategies (27%). That leaves a mere 30% of enterprises either sitting on the periphery with no strategy (8%) or in the 'consideration' phase (22%). Certainly, waiting to take action on a digital strategy could prove fatal for some organizations, according to our research.

Figure 1: Digital Transformation

Source: 451 Research's Voice of the Enterprise: Data and Analytics, Workloads and Adoption Patterns, 2019 Q. What is the current status of your organization's digital transformation initiative? Base: All Respondents



While there is no single checklist or method for what constitutes a digital transformation project, enterprise database systems continue to play a key role in these initiatives. Nearly every organization maintains a database system of record, which are often relational-based systems, holding the firm's most critical data. For that reason, any disruption to that system can lead to a material impact on the business.



But how do organizations reconcile the need to move away from older systems while simultaneously moving forward with their digital transformation initiatives? It requires a new mindset. The whole notion of a digital transformation is to change. With digital transformation initiatives, organizations must deliberately move away from outdated thinking, processes, technology and methods in favor of more modern, agile practices. For enterprises looking to transform their database systems, digitally transforming means more than getting the latest version or update; it may require a new database architecture—one that is highly scalable, fault tolerant, cloud-based and globally distributed.

SQL Isn't Back; It Never Went Away

As enterprises anticipate possible technology changes—particularly at the database level during digital transformation projects, there are some considerations to keep in mind. One concern for an enterprise moving away from a legacy relational database system is whether it can continue to leverage the SQL skills of those individuals that previously administered the system.

The Structured Query Language, which later became known as SQL, was developed by IBM in the early 1970s as a means to manipulate the data within a relational database system. Over time the SQL query language proved to be both useful and practical for database administrators, and was eventually adopted by other commercial firms that offered relational database systems. The American National Standards Institute (ANSI) adopted the SQL database query language and has maintained standards for the language since 1986.

Given that SQL has been in use for nearly 50 years, it makes sense for enterprises to have some concerns about giving it up. And while there have also been some concerns within the industry on the longevity of SQL, our research continues to show that SQL is broadly adopted and will continue to be well into the future. In fact a 451 survey of IT professionals shows that the relational data model (which requires SQL) is widely used for transactional operational workloads such as financial transactions, CRM, ERP, and sales order entry workloads.

Conversely, 10% of organizations are running non-relational data model systems, including systems categorized as NoSQL databases, which ironically often employ a type of 'SQL-like' query language to manipulate the data. Certainly, non-relational systems can handle transactional workloads, but systems employing a relational data model using SQL will continue to dominate.



Figure 2: Data Models for Transactional Workloads - Now, and in Two Years' Time

Source: 451 Research's Voice of the Enterprise: Data and Analytics, Workloads and Adoption Patterns, 2019

Q. Which of the following best describes the data model of the data platforms currently used for transactional operational workloads? What do you anticipate will be the data model of the data platforms used for these workloads two years from now? Base: All Respondents



While many enterprises prefer using a relational data model as the basis for the database system of record, it should come as no surprise that transactional workloads, which leverage the SQL query language, are the most prevalent workload for enterprises, with 62% of organizations running them on data platforms. But it should also be noted that analytical workloads, which include BI reporting, for instance, come in at 50%. These are often driven by data warehousing systems that also deploy a relational model and require SQL as the primary querying language. Hybrid workloads that blend transactional and analytical processing sit at 35% of respondents, and likewise leverage SQL for querying.



Figure 3: Workloads for Data Platforms

Source: 451 Research's Voice of the Enterprise: Data and Analytics, Workloads and Adoption Patterns, 2019 Q. Which of the following workloads are current use cases for data platforms within your organization? Base: All Respondents (n=543)



What this means for enterprises is that they do not need to leave their SQL skills behind in the event of a digital transformation project that may include full adoption of new database technologies. In fact, just the opposite is true, because in addition to transactional-based systems, many other systems that the enterprise relies on—BI reporting, ad hoc analytics and hybrid workloads, for instance—all leverage SQL querying to some extent.



Transactional Workloads Go to the Cloud

While transactional workloads are popular and will remain so into the near future, next arises the matter of where those workloads are expected to run. Historically, enterprises have been reluctant to move a transactional database to the cloud because it generally contains missioncritical data, and the organization needs to maintain certain SLA requirements without disruption.

However, few, if any, digital transformation initiatives will not include cloud computing to some degree. Therefore, cloud is expected to play a significant role for enterprises looking to adopt a digital-centric technology model. 451 Research supports this assertion as well. Currently, 70% of transactional workloads run on-premises, while 30% run in the cloud, according to a recent 451 survey of IT professionals. But the ratio is expected to change soon, with on-premises transactional workloads dropping to 48% and cloud moving up to 52% over a two-year period.

Figure 4: Location of Primary Platform for Transactional Workloads – Now, and in Two Years' Time

Source: 451 Research's Voice of the Enterprise: Data and Analytics, Workloads and Adoption Patterns, 2019 Q. Where is your organization's' primary data platform for transactional operational workloads currently located? Where will it be located two years from now?

Base: All Respondents



Transactional workloads are expected to see strong cloud adoption in the near term. But why? Certainly, enterprises are becoming more comfortable with the cloud overall, given advances in security and streamlined pricing models, for instance. But there are other reasons, too. Enterprises want to leverage the cloud as a long-term competitive advantage, and they understand that a different development and deployment approach is required.



Rethinking the Cloud, Relational Databases and Transactional Workloads

The idea of simply moving transactional workloads to the cloud as part of a digital transformation project is easier said than done. Enterprises have various options on how to move forward, including how to address existing applications, as well as how to deal with new applications:

- Lift and shift: Migrate existing on-premises applications to off-premises/cloud environments with minimal changes to the application code or business logic. While this approach does enable an organization to quickly migrate to cloud computing, it could present some challenges later on. For instance, there could be business changes such as working with different vendors, changes in the number of users and changing data requirements, all of which could have a direct impact on the application and its performance.
- **Repurchase and shift:** Replace current on-premises applications with SaaS or off-premises hosted versions of the applications. While this approach avoids some of the pitfalls of 'lifting and shifting,' it also assumes that SaaS-based applications will provide the equivalent or greater capabilities compared with existing applications, and will also be able to integrate with other systems, including legacy systems for which no SaaS application is available. This approach can work for many enterprises, but the idea of adopting new SaaS applications brings other organizational considerations.
- **Refactor and shift:** Re-architect/redesign existing applications using cloud-native frameworks and deploy them in off-premises cloud environments. This approach gets to the heart of the challenges for many enterprises. It certainly mitigates any unforeseen issues with a lift-and-shift strategy in that there are no disruptions when adopting new applications. However, while enterprises do need to invest in development and resources up-front, the long-term benefits can pay dividends given that the new application can take full advantage of cloud architecture.
- **Modernize:** Retain existing applications on-premises but move to more modern scalable, faulttolerant application and infrastructure architectures. While this approach eliminates the need to invest in development resources to refactor applications for the cloud, it does raise the issue of how legacy systems will integrate with new systems and applications as they are introduced.

But database vendors, especially those providing relational systems for transactional workloads, have likewise adapted and changed. The RDBMS has been around for approximately 50 years, which, along with SQL, has become the standard for relational systems. These relational systems have been a mainstay for enterprises that require OLTP read-write workloads that have also relied heavily on ACID guarantees for those transactions. However, changing customer and IT needs require that transactional systems be able to adapt, and those needs include adopting cloud for their data platform systems. A few specific examples are explained below.



Cloud-Native and Multicloud Adoption

As noted in the previous section, there are various methods that enterprises can choose in moving to the cloud, but there are other environmental considerations as well.

In a recent 451 Research survey, respondents were asked about their future investment plans for hybrid cloud and multicloud adoption. A combined majority of 68.2% of respondents either strongly or slightly agreed that hybrid and multicloud will be a significant investment focus going forward for their companies. That leaves a distinct minority of less than 32% of respondents whose organizations are not focusing on hybrid or multicloud in terms of investment.

Figure 5: Investment in Hybrid or Multicloud Strategy

Source: 451 Research's Voice of the Enterprise: Cloud, Hosting and Managed Services, Budgets and Outlook 2019 Q. Do you agree with the following statement: Our hybrid or multicloud strategy will be a significant investment focus in the

next 12 months. Base: All Respondents (n=393)



Enterprises also want flexibility in their choice of cloud vendor(s) by application and/or workload, in order to gain the greatest efficiencies and to protect against vendor lock-in. One approach to streamlining hybrid cloud and multicloud strategies lies in the implementation of cloud-native technologies that are developed specifically for the cloud. Such technologies include containers, service meshes, microservices, immutable infrastructure and declarative APIs, all managed by container orchestration technologies to take advantage of cloud infrastructure.



Over the last few years, these same cloud-native design principles have been brought to the world of SQL databases, which provide low-latency performance and ACID guarantees for transactions in a distributed or replicated environment. Additionally, they provide a good deal of deployment choice for enterprise customers, including:

- · Multiple datacenter or availability zone deployments
- Multiple region (aka multi-region) deployments
- Multiple cloud (aka multicloud) platform deployments.

Scaling in the Cloud

Another significant enterprise need centers on adapting to the distributed nature of cloud architecture and the ability for transactional systems to better accommodate scaling—traditionally an Achilles' heel for relational systems.

For relational databases, there are two fundamental ways to scale: scale up or scale out. Scaling up refers to increasing the capacity of an existing system by adding memory, storage or computing power. The idea is to increase the capacity of the system, which in turn improves performance and availability, for instance. However, modern applications may require greater scaling beyond a scale-up approach. Scaling out, on the other hand, involves adding servers to accommodate additional data or users, as a means to spread out the load. Scaling out has its advantages; the most obvious is that it addresses excess load on the system, allowing the overall database to expand and shrink as needed. For instance, there is the option to scale down in a scale-out architecture. If load decreases for some time, nodes can be easily removed and repurposed.

It is interesting to note how important scaling becomes as enterprises consider their digital transformation initiatives. 451 Research survey data indicates that while 70% of IT professionals prefer to deploy a scale-up architecture for their transactional workloads today, this figure is expected to drop to just 59% in two years, with 41% indicating a preference for scale-out architecture in their future data platform deployments.



Benefits of Data Systems

Enterprises embark on digital transformation initiatives for a myriad of reasons, but always have certain expected benefits they are looking to obtain. In a recent 451 Research survey, IT professionals were asked about the benefits they were looking for from their database systems. More than half (55%) of respondents said they are looking specifically for operational efficiency from their data platforms. Improving customer experience came in second at 45% of respondents. All other benefits—such as gaining a competitive advantage, increasing sales, lowering costs and reducing business mistakes—were cited by less than 30% of respondents.

Figure 6: Benefits of Data and Analytics

Source: 451 Research's Voice of the Enterprise: Data and Analytics, Workloads and Adoption Patterns, 2019 Q. What are the most significant benefits your organization is realizing from its use of Data and Analytics? Base: All Respondents (n=585)



While operational efficiency can speak to lowering costs and reducing business errors—both choices provided in the survey—what organizations are looking for is a direct correlation between inputs and outputs. In other words, organizations are looking to maximize their data platform investment (input) with expected benefits (output) on the back end. In many ways, this makes perfect sense for why organizations may be looking to migrate their transactional workloads to the cloud as part of a digital transformation initiative.



Enterprises are not interested in just migrating to the cloud for the sake of keeping up with the competition, important as that may be. Ultimately enterprises are interested in achieving operational efficiencies when migrating to the cloud. As such, operational efficiencies come when enterprises can deploy their data platforms in the cloud in order to fully leverage elastic infrastructure as a means to provide massive scaling, distributed transactions, high availability, and the ability to deploy across multiple regions, datacenters and cloud platforms.

Recommendations

- Digital transformation initiatives can be a significant undertaking for many organizations, and will impact their systems of record. Avoiding disruption to these systems when newer, more advanced technologies exist could have long-term negative consequences.
- Enterprises need not assume that a change in their relational database requires giving up on SQL expertise among their most skilled workers. SQL is expected to remain a required IT skill well into the future, and can be used for newer relational database technologies.
- A cloud computing strategy should be a core component of an organization's digital transformation journey because research indicates that while transactional workloads will remain strong, they will eventually reside in the cloud.
- Organizations should realize that transactional workloads running in the cloud can provide several benefits, including massive scaling, distributed transactions, high availability, and the ability to deploy across multiple regions, datacenters and cloud platforms.
- Enterprises will find that pairing cloud-native infrastructure with a cloud-native database represents the best path forward to ultimately driving organizational efficiency.



Yugabyte is the company behind <u>YugabyteDB</u>, the open source highperformance distributed SQL database for building global, internet-scale applications. YugabyteDB serves transactional applications with SQL query flexibility and cloud native agility, thus allowing enterprises to focus on business growth instead of complex data infrastructure management. Customers can choose from two DBaaS solutions: <u>Yugabyte Platform</u>, a management platform for building private YugabyteDB-as-a-Service offerings on any cloud or Kubernetes environment, and <u>Yugabyte Cloud</u>, a fully managed pay-as-you-go cloud database available on major public cloud platforms. Learn more at <u>Yugabyte.com</u> and schedule a free demo.



About 451 Research

451 Research is a leading information technology research and advisory company focusing on technology innovation and market disruption. More than 100 analysts and consultants provide essential insight to more than 1,000 client organizations globally through a combination of syndicated research and data, advisory and go-to-market services, and live events. Founded in 2000, 451 Research is a part of S&P Global Market Intelligence.

© 2020 451 Research, LLC and/or its Affiliates. All Rights Reserved. Reproduction and distribution of this publication, in whole or in part, in any form without prior written permission is forbidden. The terms of use regarding distribution, both internally and externally, shall be governed by the terms laid out in your Service Agreement with 451 Research and/or its Affiliates. The information contained herein has been obtained from sources believed to be reliable. 451 Research disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although 451 Research may discuss legal issues related to the information technology business, 451 Research does not provide legal advice or services and their research should not be construed or used as such.

451 Research shall have no liability for errors, omissions or inadequacies in the information contained herein or for interpretations thereof. The reader assumes sole responsibility for the selection of these materials to achieve its intended results. The opinions expressed herein are subject to change without notice.



NEW YORK 55 Water Street New York, NY 10041 +1 212 505 3030

SAN FRANCISCO One California Street, 31st Floor San Francisco, CA 94111 +1 212 505 3030



LONDON

20 Canada Square Canary Wharf London E14 5LH, UK +44 (0) 203 929 5700

BOSTON 75-101 Federal Street Boston, MA 02110 +1 617 598 7200

