IDC InfoBrief
June 2022

Competing in the Digital-First Economy

New Engines for Enterprise Data Workloads

IDC Doc. #AP241313IB

IDC

Sponsored by PingCAP

65% of global GDP will be digitalized in 2022

The power of digital transformation (DX) has never been felt so strongly as in the wake of the COVID-19 pandemic. It is with the help of digital products and services that people stay connected, communities continue functioning, and businesses thrive despite the many inconveniences and disruptions.

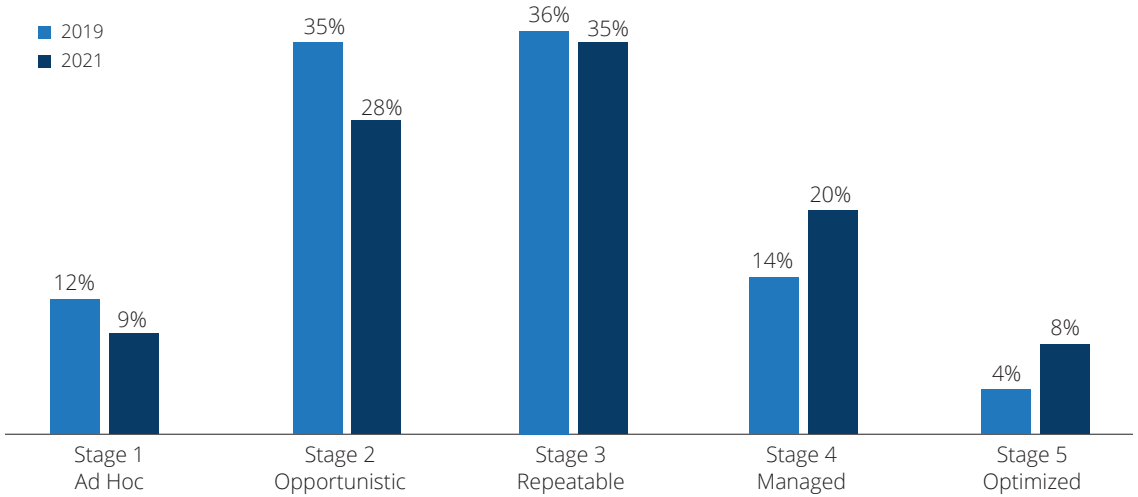
In Asia/Pacific (including Japan) (APJ)

28% of organizations are in the advanced stages (Stages 4 & 5) of DX.

DX technology spending will grow **2x faster** than overall IT spending.

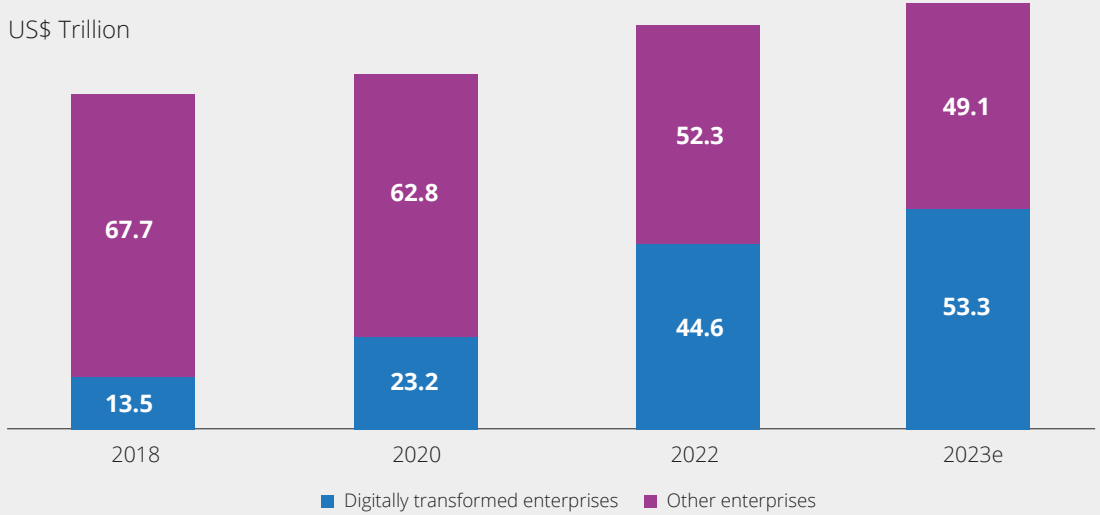
55% will invest in "grow" or "transform" initiatives.

APJ DX maturity 2019 (pre-COVID-19) vs. 2021 (post-COVID-19 outbreak)



A digital-first world is underpinned by a digital-first economy where GDPs see the biggest contribution from digitalized products and services.

Worldwide nominal GDP driven by digitally transformed vs. other enterprises



IDC predicts that the economy remains on course to its digital destiny with 65% of global GDP digitalized by 2022 and will drive over US$6.8 trillion of direct DX investments from 2020 to 2023.

Source: IDC's APJ Business and Technology Leaders' DX Impact Survey 2021 (APJ N = 513); IDC Future Enterprise Awards 2021; and IDC FutureScape: Worldwide Digital Transformation 2021 Predictions, Doc #US46880818, October 2020

Future Enterprise guide: Surviving and thriving in the digital-first economy

The Future Enterprise is IDC's vision of how companies must organize and invest to participate in the digital-first markets. It is a framework comprising nine distinctive and interrelated dimensions, four of which are of significance to business leaders – customer experience, operations, work, and industry.

The Future Enterprise

FUTURE OF
Customer Experience

Always-on and empathetic customer engagement

through their preferred channels

FUTURE OF
Operations

Simplified core

to orchestrate critical operations to become responsive, preemptive, and adaptive

FUTURE OF
Work

Task execution by hybrid workforce,

leveraging gig-economy workers and digital workforce (bots)

FUTURE OF
Industry Ecosystems

Resilient value chain

through shared data and shared expertise

All enabled by trusted data workloads that unify and scale success across geo-locations

Four key characteristics of Future Enterprise data workloads

The core competencies of Future Enterprises are enabled by trusted data workloads that handle **massive volumes of highly concurrent, real-time transactions** with customers, workforce, and business partners at any time, from anywhere, in a **highly scalable and adaptive** manner.

Scalable

Up and out, in a cost-effective manner

Anytime, anywhere

The global datasphere of structured data is predicted to grow at a compound annual growth rate (CAGR) of 42% in the forecast period of 2020–2025. Organizations will therefore increasingly need to manage data workloads cost effectively.

Real-time

Insights as transaction commits

Augmented as business runs

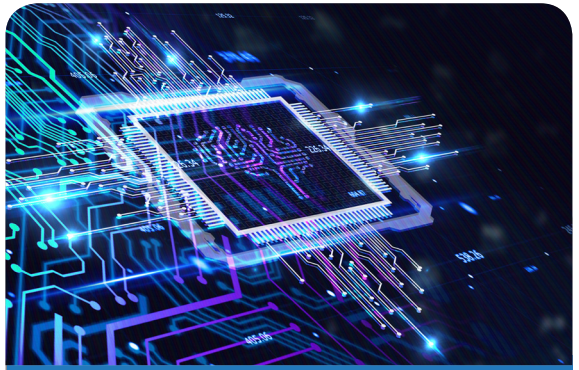
Intelligent product recommendation, service routing, insight generation, content curation, and resource optimization are becoming default features, to be delivered at the speed of business.

Adaptive

As apps and devices proliferate

Anticipate, adapt, agile

DX is a means for organizations to achieve agile innovation. It is a continuous pursuit in response to fast-paced market dynamics and ongoing disruptions.



Resilient

For mission-critical workloads

Always on, always available

With digital channels established, it is common for large organizations to have mission-critical transactions reaching tens of thousands, even millions, of queries per second (QPS).

Source: Worldwide Global DataSphere and Global StorageSphere Structured and Unstructured Data Forecast, 2021–2025, Doc #US47998321, July 2021

The rewards of doing it right in the digital-first economy

Leading organizations are already reaping tremendous rewards. But this is no easy feat. When things go wrong, organizations face severe consequences and significant penalties.

Severe consequences – when services become unavailable

A leading bank in Southeast Asia
15 hours of service disruption affected around 2.5 million users in Singapore in November 2021.

A leading ride-hailing service provider in Southeast Asia
4 hours of disruption affected around 650 million users in November 2021.

A leading global social media platform
6 hours of disruption affected billions of users worldwide in October 2021, causing the company's shares to go down 5% on the day of the incident.

> "While our business boomed, the soaring data size became a nightmare for us. Every day, our billing applications processed over 3.6 billion requests, generated over 5 terabytes of log data, and wrote dozens of gigabytes of data into the database. This exposed the bottlenecks of our standalone MySQL architecture in scalability, data isolation, performance, and other areas."
>
> **NetEase Games**

Rewards – when a new experience goes viral

From China to ~ 150 international markets

TikTok

AI-powered short-form video social media platform

- Gained 30 million international users in less than 3 months in 2018
- 2020 revenue growth (year-over-year): 111%
- 2 billion active users in 75 languages as of 2021

Riding on the work-from-home opportunity

Zoom

Videoconferencing platform for work and play

- Mobile app downloads increased 7x within 1 month (March to April 2020)
- Estimated number of daily active users: 400 million
- 2020 revenue growth (year-over-year): 869%

Scalable | Real-time | Adaptive | Resilient

Rising challenges in digitally transformed enterprise data workloads

CHALLENGE 1

Lack of adaptiveness of data pipelines

Data fragmentation and movement

needed for continuous innovation

The reality: Heterogeneity of the data ecosystem brought by:

Application innovation

(Hybrid) infrastructure upgrade

Device proliferation

The result: New data sources and new data type mixes to feed into new data pipelines

Dozens of new data source requests per month

Multiple types of data in each data request

CHALLENGE 2

Lack of real-time capabilities

Real-time complex query

needed for improved experience

Increasingly delivered in real-time, and with intelligence embedded:

Dozens of new delivery requests per month

Over 50% of AI use cases require real-time inferencing

CHALLENGE 3

Lack of operational resilience

Global scale out

needed for market expansion

To support resilient global operations of massive volume with high concurrency, and compliance with data residency requirements:

Operational resilience

Tens of thousands to millions of QPS

⚠ Data residency and compliance

In 2021, a company was fined US$877 million – the largest fine so far – for violation of data privacy and protection regulations.

Dealing with data fragmentation: Cloud-native and open source for adaptive innovation

CHALLENGE 1

Lack of adaptiveness of data pipelines

Data fragmentation and movement

Cloud-native

Ensures agile and continuous development and deployment

The deployment of enterprise applications in APJ is highly hybrid – a mix of public cloud, industry cloud, private cloud, and on-edge and on-premises deployment is the norm. **Cloud-native is the only way to ensure agile deployment anywhere**.

Already nearly **43% of APJ organizations favor cloud-native solutions** when committing resources for DX, to make it possible to use a wide range of platforms from multiple locations.

> "We deployed databases independently for different game products. Data was scattered on data silos (that is, isolated islands of data), which made it difficult to gain insights from data analysis. When performing cross-product calculations, we needed to maintain multiple heterogeneous data sources, and the methods to access these sources were complicated."
>
> **A large gaming company based in China**

Open-source as glue

To bring together a heterogeneous data ecosystem adaptively

Data fragmentation is inevitable and persistent – new apps and new Internet of Things (IoT) devices are continuously being added to enterprises' IT landscapes. This has created concrete demand for productivity tools to discover and update metadata and automate data pipelines for analytical use cases. **Open-source products are the friendliest for such productivity tools to work with agility and flexibility**.

Open-source for innovation

Customize by organization's own innovation agenda, and avoid vendor lock-in

Open-source is becoming the center of gravity in the world of software development. It effectively **leverages the collective intelligence of a very large developer community to accelerate industry ecosystem level innovation**. With the increasing size of such communities (with year-on-year growth of active GitHub contributors in 2020 going beyond 21%), more disruptions are expected in the world of enterprise solutions.

APJ respondents show between **4 and 5 times the preponderance** of other geographies for using open-source software as a tool to prevent future lock-in with any single provider's foundational cloud services portfolio.

Source: IDC Market Perspective, Considering Role of Open Source Software on Cloud Platforms for Digital Transformation Projects, Doc #US48320921, November 2021

Achieving real-time intelligence: HTAP for a unified world of OLTP and OLAP

CHALLENGE 2

Lack of real-time capabilities

Real-time complex queries

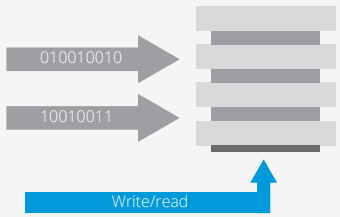
Hybrid transaction and analytic processing (HTAP)

In a digitalized world of abundant choices, intelligent curation and action recommendation are becoming must-have features for consumer-facing services. In response to this, data capture and movement technology in Asia/Pacific (excluding Japan) (APEJ) will increasingly be used to enable real-time simulation, optimization, and recommendation capabilities. **Databases that can support HTAP in the same database session are becoming a big deal for mission-critical applications**.

> To follow regulations, the calculation for suspicious patterns and risk rating must be completed the day after the payment is made. This requirement poses a challenge to the processing time of our system. Previously, a batch processing task would take 7–8 hours, and the overall task would last for 15 hours or more each day. As the data volume surges, our application system has a greater risk of missing the performance target.
>
> **Mobile payment subsidiary of a leading global telco provider**

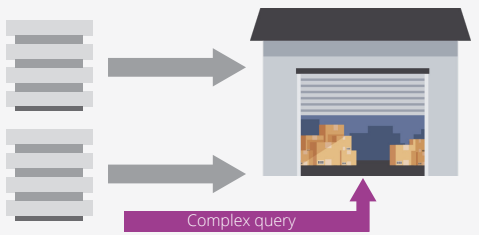
Where are you in the journey to HTAP?

Online transaction processing (OLTP)



Dylan bought a tennis racket for US$100 at bestsportsgears.com

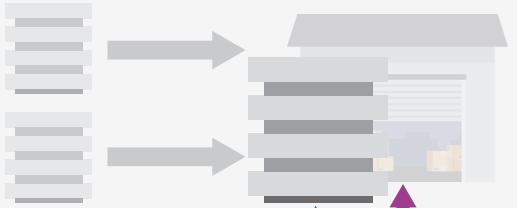
Online analytical processing (OLAP)



Dylan's next likely purchase includes gym membership, dry-fit shirts, and bicycles, etc.

HTAP

Supports complex analytic query and transaction processing against a single database



Dylan bought a tennis racket for US$100 at bestsportsgears.com

His next likely purchase includes gym membership, dry-fit shirts, bicycles, etc.

Scaling out for resilience: Highly distributed, consistent, and performant

CHALLENGE 3

Lack of operational resilience

Global scale out

Distributed for resilience

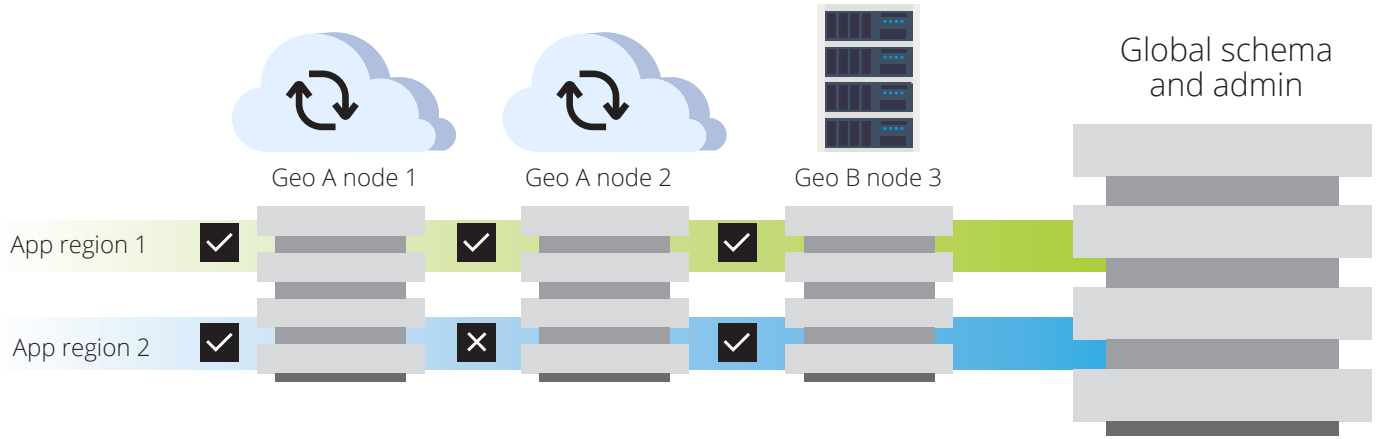
Distributed databases are needed for resilient operations against potential node failures, as well as compliance requirement for data residency. Geographical diversity of distributed databases allows data to be stored with replicas available on different physical nodes at different locations, by residing on distributed and often hybrid infrastructure – independent servers and datacenters, provisioned on public or private cloud.

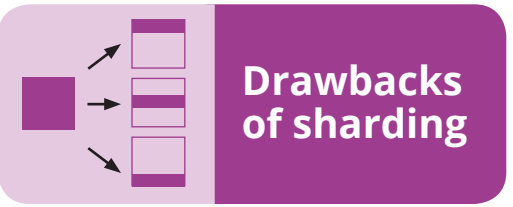
Strong consistency and high performance

As is with all distributed systems, consistency across all nodes is required regardless of change initiator. Consistency is subject to network performance and the design of synchronization operations, as well as rollover mechanism when failure occurs, etc. **Ensuring consistency becomes more challenging (1) when the number of nodes increases (higher scalability), and (2) the system throughput increases (higher performance)**.

> As data size increased, storage and computing issues became more and more severe. Unfortunately, we couldn't solve this problem by simply upgrading hardware – in any case, that option was expensive. Therefore, we decided to look for a new solution.
>
> **A leading express delivery company**

Improving performance, scalability, and consistency of a highly distributed system at the same time is challenging



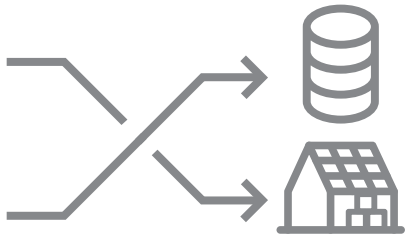


Sharding is a database partition method to help applications horizontally scale out. However, application level sharding introduces complexity and rigidity into the database architecture, **increasing the risk of data loss, table corruption, and creating lock-ins for technical support**.

Characterizing the new engines of future enterprise data workloads

Cloud native and open source

To tackle data fragmentation



HTAP

To facilitate real-time complex queries

Distributed with strong consistency

Global scale out with resilience

The advantages of DBaaS

For cost effectiveness and skill leverage

The cloud database market grows at a five-year CAGR of 36%, 4.5x that of the on-premises market (APEJ, 2020-2025)

Scale as you need: Scalability in the digital-first world is easier said than done. Peak demand traffic can be totally unplanned; businesses need to make preparations without the operational overhead to maintain the distributed system.

Peace of mind: Managed cloud database services, also called database as a service (DBaaS), take most of the operational responsibility away from the IT staff and place it in the hands of the service providers.

Subject to the test of real-world use cases and demand surges

HTAP capabilities are varied for different analytical use cases – depending on how complex the queries are, whether machine learning is required, etc.

Quality of open-source projects depends on the size, and the activity level of the contributor's community.

Consistency of distributed systems is subject to the battle test of massive volume transactions and demand surges in international markets covering different deployment types.

High scalability for fast growing digital-native businesses

	Square	Databricks	U-Next	Ninja Van
Company profile	▪ Leading global payment app	▪ Founded by the creators of Apache Spark ▪ Develops industry leading cloud data lakehouse platform that allows analytical queries of structured, semi-structured, and unstructured data	▪ One of the largest media companies providing subscription streaming service in Japan	▪ A leading logistics company in Southeast Asia headquartered in Singapore ▪ Operates in 6 Southeast Asian countries
Workload profile	▪ ~2.5 million queries per second	▪ Tens of terabytes of data across AWS, GCP, and Azure in close to 100 regions	▪ Tables grew by 300 million rows every month	▪ Query tables as large as 10TB

> "Our belief is that TiDB will support the rapid growth of Databricks as a business and allow us to grow faster in the future."

Reynold Xin, Co-Founder and Chief Architect of Databricks

Resilient operations for large enterprises

	Xiaomi	Bank of Beijing	UnionPay	Ping An Insurance
Company profile	▪ Leading consumer electronics and software company ▪ 4th largest smartphone manufacturer in the world	▪ An urban commercial bank based in China	▪ The only interbank payment service provider in China	▪ Leading life insurance company in China ▪ More than 200 million registered users
Workload profile	▪ 100 million read and write queries daily (to the MIUI operating system)	▪ 7,500 write queries per second	▪ ~240,000 transactions per second	▪ Up to RMB100 billion payment volume in a single day ▪ ~10 million daily active users during peak time

> "NewSQL databases can be scaled on demand, dynamically adapting the entire system's performance to meet uncertain business needs. The storage and query of massive structured data are no longer a headache, greatly improving the efficiency of our application development. TiDB performs well in steady-state business scenarios. It is also an ideal choice in agile business scenarios."
>
> **Ping An Jin Guanjia Application Development Team**

Real-time analytics for business insights

	Xiaohongshu	ZTO Express	Zhihu	Streak
Company profile	▪ China-based social media company (similar to Instagram) ▪ 160 million active users	▪ One of the largest leading express delivery companies in the world based in China	▪ China-based Q&A platform (similar to Quora) ▪ 220 million registered users	▪ A Google Chrome extension that creates a customer relationship management (CRM) dashboard ▪ ~6,000 customers worldwide
Workload profile	▪ 10 billion read queries per day	▪ 17 billion parcels delivered in 2020 ▪ 1.4 billion monitoring items added per day	▪ 40,000 records written per second	▪ 35TB of email metadata across the total user base

"TiDB is enabling Streak users to share emails and metadata effortlessly, and is enabling our developers to move faster than ever before."

Fred Wulff, VP of Engineering

Transform your business to become future ready

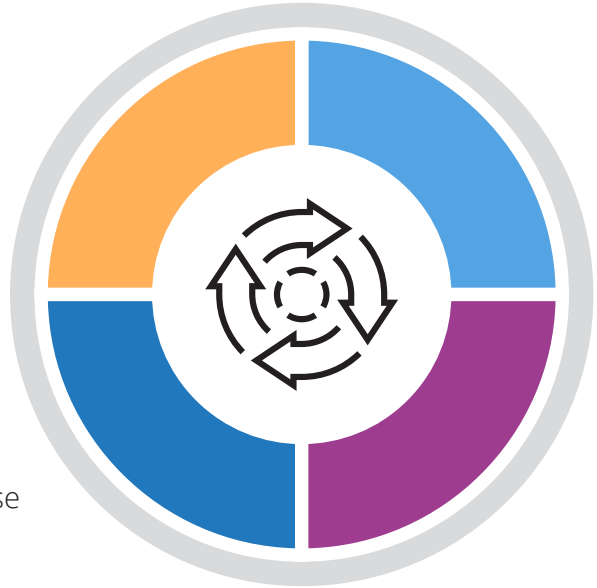
Are you up to the challenge of being a Future Enterprise?

FUTURE OF
Customer Experience

Always-on customers with their own preferred channels, and very limited attention span

FUTURE OF
Work

Hybrid workforce, leveraging gig-economy workers and digital workforce (bots)



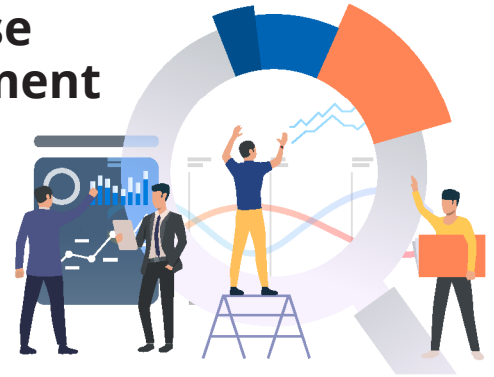
FUTURE OF
Operations

Agile collaboration with business partners with close to real-time transparency

FUTURE OF
Industry Ecosystems

Resilient industry ecosystem enabled by shared data and expertise to thrive despite disruptions

How you can transform your enterprise data management



Scalable

Achieve market expansion through unified and scalable online experience and working model, in a cost-effective manner.

Real-time

Deliver products, services, and experience augmentation with the speed of business.

Adaptive

Commit resources for continuous business model innovation to adapt to market dynamics and best leverage technology advancement.

Resilient

Modernize mission-critical legacy applications to engage with a huge number of customers, workforce, and business partners.

Manage enterprise data workload by specialization and unification

Technologies are becoming more specialized when interfacing with data, and at the same time, more unified when interfacing with data workers.

Organizations need to:

Understand where they are in the evolving picture of enterprise data management.

Architect with the distinct traits of new enterprise data engines in mind – cloud native and open source, HTAP, distributed with strong consistency, and cloud DBaaS.

Enterprise data management

Operational data management

Large volume and operational (**timeliness** matters)

Relational data management platforms

RDBMS

Specialized data management platforms

Stream processing

NoSQL DBMS (such as Graph)

HTAP

Analytic data management

Exploratory and consolidated (**comprehensiveness** matters)

Data fabric

Data warehouse

Data lake

Data lakehouse

Data mart

Data supermarket

(Front-desk) Data-driven business applications and (middle-desk) analytic pipelines

Operation

Reporting | *Exploration*

Where unification emerges

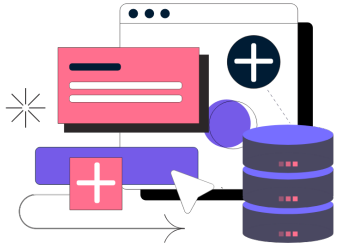
Message from the sponsor

Power your digital business with TiDB Cloud

A cloud-native, distributed SQL database built for real-time analytics in a fully managed service

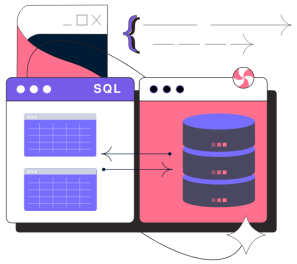
TiDB Cloud

Attributes



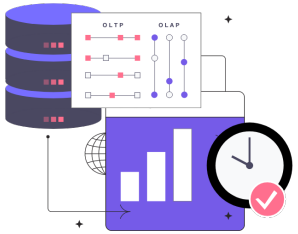
Fast scaling

Elastically and transparently **scale to hundreds of nodes** for critical workloads without changing business logic.



Simple convergence

Use what you know about SQL – maintain your **relational model and global ACID transactions** while coping with your hybrid workloads at ease.



Real-time analytics

Equipped with **built-in high performance analytics engine** to analyze operational data without using an ETL.

Key features

Horizontal scalability

Scale out to hundreds of nodes while maintaining ACID transactions. No need to bother with sharding or facing downtime. Ensure data accuracy at scale – even for simultaneous updates to the same data source.

MySQL compatibility

Increase productivity and shorten time-to-market for your applications with TiDB's MySQL compatibility. Easily migrate data from existing MySQL instances without the need to rewrite code.

High availability

Naturally high availability by design. TiDB's Auto-failover and self-healing ensure business continuity, regardless of hardware failure, network partition or data center loss.

Operational insights

No ETL and business interference. With a built-in analytics engine, TiDB Cloud runs consistent analytical queries on current data without disturbing mission-critical tasks.

Fully managed

Use it instead of operating it. TiDB Cloud provides fully managed services on AWS and GCP, with enterprise grade security features like data encryption and auditing. Certified by SOC 2 Type 2, ISO 27001:2013 and fully compliant with GDPR.

Battle tested

More than 2,000 adopters globally. Trusted and verified by innovation leaders in banking, fintech, ecommerce, and gaming.

PingCAP

Apply now for TiDB Cloud PoC with expert technical support!

Learn more at https://en.pingcap.com/tidb-cloud

IDC Asia/Pacific
83 Clemenceau Avenue
#17-01 UE Square, West Wing
Singapore 239920
T 65.6226.0330

idc.com

@idc

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC Custom Solutions

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.

Copyright 2022 IDC. Reproduction is forbidden unless authorized. All rights reserved.

Permissions: External Publication of IDC Information and Data
Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager.
A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason. Email: ap_permissions@idc.com

IDC Doc. #AP241313IB