Databases in the Cloud

There’s nothing like response, redundancy and optimal management

White Paper

IONOS

Contents

1 Introduction

The digital transformation is gaining momentum at a rate of knots. In a process involving the constant development of digital technologies, companies of all shapes and sizes are looking to digitalise their existing business models and develop new ones. According to the business consultancy etventure, the coronavirus pandemic has given fresh impetus to the developments after the digital transformation had declined within large German companies for the first time in 2019[1].

One of the key drivers behind the digital transformation is the continuously expanding volume of data. Many companies are still encountering difficulties when it comes to consistently realising their potential – according to the etventure survey cited above, the greatest obstacle is the lack of qualified IT experts. According to the market researchers at IDC, 463 billion exabytes of data will be generated every day by the year 2025[2]. And it is particularly worth highlighting the exponential acceleration of this growth: More data will be produced over the next three years than in the past 30 years combined. In fact, over the next five years, our digital ecosystems will generate three times more data than the amount produced in the last five years[3]. One of the main drivers in this field is the Internet of Things (IoT).

In today's data economy, this data has to be constantly aggregated and made available for organisations to use to their advantage and run their business models. Data was once described as "the new oil" – and this statement now rings truer than ever before. At the same time, new data sources are constantly being created in the age of Industry 4.0 and the Internet of Things, which is similarly accelerating the development outlined above.

When it comes to collecting data and making it available, modern database infrastructures are some of the central elements. They must be capable of storing the exponentially growing volumes of data while enabling real-time analyses and evaluations that are critical to many digital business models. A clear trend has emerged in recent years with regard to storage and computing services in general and with regard to databases in particular: A growing amount of data has been moved to the cloud. The advantages of the cloud in terms of flexibility, efficiency, scalability and cost optimisation enable companies to easily and securely run the high-performance databases needed to ensure their current and future success.

This can even be requested as a managed service in a lean infrastructure to reduce the unnecessary and unwanted costs incurred by customers for the operation and management of their databases and to free up resources for more creative challenges.

There are various ways of running a database in the cloud – from simple virtualisation and container solutions to the ultimate cloud-based model: Database-as-a-Service (DBaaS).

---

[1] etventure (2019): "Studie Digitale Transformation: Die Zukunftsfähigkeit deutscher Unternehmen", accessible online at: https://service.etventure.de/digitale-transformation-2019-registrierung.

[2] Quoted from World Economic Forum (2019): "How much data is generated each day", accessible online at: https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/.

[3] IDC (2020): "IDC's Global DataSphere Forecast Shows Continued Steady Growth in the Creation and Consumption of Data", accessible online at: https://www.idc.com/getdoc.jsp?containerId=prUS46286020.

2 What is Database-as-a-Service?

DBaaS is a cloud-based approach to managing and storing data. As part of the platform as a service (PaaS) family, the cloud provider not only offers the necessary hardware and network infrastructure, but also database infrastructures and database management systems. If we think about the distinction that is usually made between infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS), DBaaS is positioned somewhere between PaaS and SaaS.

DBaaS uses the platform's underlying infrastructure, which is run securely at the provider's data centre, and utilises its computing power, storage space and network capacity. The data centre is either connected to the Internet via high bandwidths or can be controlled directly from the user's network via dedicated connections. Users can access their databases via special interfaces to save, view and edit data. They do not need their own database infrastructure.

When DBaaS is requested as a managed service, the provider is not only responsible for hosting the database – it also takes care of administration, operation, maintenance and continuous updates. In addition, the provider manages and automates the following functions and workflows:

- **Rapid provisioning:** The cloud provider automatically deploys the necessary resources.
- **Scaling:** The cloud provider makes sure an increasing volume of data and work-loads can be processed.
- **High availability:** The cloud provider installs redundant and distributed systems and performs data back-ups to make sure the database environment is robust enough to continue providing the required services in the event of malfunctions and other problems.

- **Failover:** The cloud provider establishes automated processes with replacement systems to protect the database systems against failure and ensure maximum availability.
- **Back-up:** The cloud provider copies the database environment and the data stored there so that it can be restored in the event of data loss.
- **Geo-replication:** The cloud provider duplicates data at geographically distant locations so that back-up copies are available in the event of a disaster.
- **Monitoring and alerts:** The cloud provider uses tools to monitor and optimise the database infrastructures and systems.
- **Support:** The cloud provider offers technical support if users have any questions or issues.

As part of DBaaS, users can opt for both relational and non-relational databases. Many companies are now using applications that require different database technologies. DBaaS solutions enable companies to simply and reliably use the high-performance technology they need with a high degree of data security – without the need for extensive expertise in the relevant fields. As the services are billed according to usage, customers only pay for the resources they actually need. In addition, DBaaS often includes native tools for analysing the data stored in the databases.

3 How is DBaaS different to traditional database hosting?

DBaaS isn't the only way databases can be moved to the cloud; there are various cloud-based models which differ in particular with regard to the effort required from users when it comes to running and managing the database.

3.1 Self-managed databases with virtualisation

Self-managed databases with virtualisation: The databases are not operated on premise in the user's infrastructure, but via a cloud provider using IaaS in a virtualised instance.

> ***Advantages:*** *Self-managed databases in the cloud offer customers a high degree of flexibility and control. Any parameters can be changed or adapted at any time if required; at the same time, customers can watch over their data, such as in a private cloud environment. That's why companies often use this option for highly sensitive data. Their external costs usually remain manageable, because they simply outsource the infrastructure.*

Disadvantages: *It takes a lot of resources to manage databases. As they have to be operated, serviced, maintained, optimised and otherwise handled by specially employed personnel, there are high fixed costs and the flexibility gained elsewhere is limited. In addition, self-managed databases usually have limited scaling options, so the services soon reach their limits during peak loads and important business processes may come to a standstill as a result. Another disadvantage is the fact that companies have to take care of the security of their own systems, including regular maintenance, updates and patches.*

Apart from hosting, no other services are requested from the cloud provider.

3.2 Databases with containers or Kubernetes

- **Databases with containers or Kubernetes:** These are similar to self-managed databases but run on a containerised infrastructure. However, customers first have to check whether the required workloads are even suitable for Kubernetes.

Advantages: *By using Kubernetes, data controllers can take advantage of the increased automation possibilities offered by containers. At the same time, containerised databases enable greater portability for pending migration processes.*

Disadvantages: *Databases running on Kubernetes are more prone to errors than self-managed databases due to the transient nature of pods. In addition, users have to spend more on the management and operation of their databases compared to DBaaS, because they have to take care of their own operations.*

3.3 Database-as-a-Service

- **Database-as-a-Service:** In addition to hosting the database, the provider takes care of the entire configuration, operation, scaling and necessary back-ups as a managed service. This means users do not have to spend any resources on these services – they can focus on more useful tasks (e.g. app development) after their database has been configured.

3.4 Managed service from a third-party provider

- **Managed service from a third-party provider:** This is where companies get everything from one provider. The provider acquires and administers the necessary licences, installs and configures the database, monitors and backs up the data, and even offers comprehensive advice on the appropriate solutions before the project gets under way. In contrast to DBaaS, organisations can access the necessary, regular-sized databases directly via self-service portals – without having to waste their time on unnecessary steps.

4 Why in the cloud? What are the market volumes and potential?

Back in 2019, the analysts at Gartner published a report entitled "The Future of the Database Management System (DBMS) Market is Cloud"[4]. In the year preceding its publication, cloud-based services accounted for nearly a quarter of the DBMS market. Their forecast has remained accurate to this day, as 75% of databases around the world will run in the cloud by 2022.

This trend has been confirmed by the market researchers at Unisphere Research[5].

Almost half of the companies included in a 2019 survey stated that they hired cloud providers to implement new database projects – although some of them favoured a hybrid approach where parts of their databases were run on premise and other parts were administered as a managed service. However, an overall trend towards the cloud can also be observed here: While 60% of the respondents favoured their own data centres for the configuration of new databases in 2019, Unisphere Research expects only 26% of companies to take this approach by 2022.

The shifting of databases to the cloud is even one of the biggest trends in more recent studies. In a survey conducted by the database provider MariaDB in autumn 2020, 47% of the companies in Germany stated they would be moving to the cloud later that year[6]. In addition to the increasing automation of databases (54%), the implementation of multi-cloud environments (37%) and the relocation of analytical data to the cloud (37%), these trends are a direct reflection of the high demand for cloud-based databases.

---

[4] Gartner (2019): "The Future of the DBMS Market is Cloud!", accessible online at: https://blogs.gartner.com/adam-ronthal/2019/06/23/future-database-management-systems-cloud/.

[5] Unisphere Research (2019): "2019 IOUG Data Environment Expansion Survey", accessible online at: https://www.dbta.com/ DBTA-Downloads/ResearchReports/2019-IOUG-Data-Environment-Expansion-Survey-8595.aspx

[6] Computerwelt (2020): "Datenbank-Trends: Automatisierung und Cloud-Migration", accessible online at: https://computerwelt.at/news/datenbank-trends-automatisierung-und-cloud-migration/.

DBaaS has become one of the fastest growing cloud services in the world. According to Research and Markets, the global DBaaS market will grow to over $ 320 billion by 2025[7]. Companies are looking to tap into the numerous advantages of DBaaS, which make cloudification as a service a highly attractive prospect from both a technical and business perspective.

5 What are the advantages of cloudification services?

- **High flexibility and elasticity:**
  Databases often have to deal with highly dynamic workloads and many companies find it difficult to reliably forecast such information. DBaaS makes it quick and easy for companies to adjust the required resources, especially when the conditions change, such as by flexibly adding or removing additional server capacities as required.

- **Good scalability:**
  As DBaaS is more flexible and elastic, it is a highly scalable service. The required services can be quickly adapted to changing requirements (e.g. seasonal peaks in e-commerce, upcoming quarterly reports).
  The services can usually be changed by users in real time via the DBaaS management interface.

- **Low error rate:**
  DBaaS is characterised by a high degree of automation. By automating standard processes such as the installation of updates and patches, the error rate is reduced at the same time.

- **High productivity:**
  The automation possibilities offered by DBaaS lead to greater productivity. As customers no longer have to inspect, maintain and manage their databases, they can focus on more creative tasks such as developing digital business models or consulting their departments.

- **High cost efficiency:**
  Companies can always request the exact service they need to meet their changing application requirements. In this way, the rigid CAPEX spent on an infrastructure prepared for all eventualities make way for flexible OPEX. Any additional capacities that are required can be booked quickly, flexibly and for a limited period of time without causing any business interruptions. TCO calculators usually enable users to estimate the costs incurred for special applications and the required computing in advance, which further simplifies the decision-making process.

[7] Research and Markets (2019): "Global Cloud Database & DBaaS Market Report", 2019: Industry Forecast to 2025", accessible online at: https://www.businesswire.com/news/home/20190809005257/en/Global-Cloud-Database-DBaaS-Market-Report-2019.

- **High availability:**
  The DBaaS hardware is located at professional data centres operated by the cloud provider, which are very well protected against failures. The providers also have many years of experience and extensive expertise in the operation of hardware and database services. The extensive automation possibilities offered by DBaaS also ensure high availability: As updates are generally installed in real time, the database services can continue to be used in parallel without restrictions. Most providers offer 99.99% availability or more. If companies want even greater availability, they have to reach out to DBaaS providers whose infrastructure is distributed across different availability zones.

- **Managed service without your own infrastructure:**
  Companies do not need their own dedicated servers or other infrastructure to request DBaaS. This means the performance of their databases does not depend on the hardware resources at their disposal. After all, this infrastructure is supplied by the provider alongside the necessary database software. The provider automatically installs safety and OS updates (e.g. patches). In addition, users are guaranteed a certain degree of availability and database performance through service-level agreements (SLAs). This minimises the need for in-house expertise and personnel capable of looking after their databases. It also strengthens reliability and predictability.Habe

- **Less administration and quick deployment:**
  Automated workflows and one-click operations simplify the database management process enormously. When using DBaaS, distributed silo databases can be consolidated under one roof upon request. The centralised environment saves time and precious resources. Users only have to focus on the applications they want – the managed service provider takes care of runtimes, middleware, operating systems, virtualisation, servers, storage and networks.

- **High data security:**
  This is usually integrated in cloud-based infrastructures, consistently implemented across the entire platform and continuously improved through white hat penetration tests to identify potential vulnerabilities. DBaaS providers usually have sophisticated security and encryption technology. This includes the encryption of data in transport and in the database itself, as well as monitoring and key management. Other frequently used security services include dynamic data masking and the use of tokens.

- **High speed:**
  Depending on the requested product, companies can set up a database with a DBaaS provider in a matter of minutes – that's a huge difference compared to the lengthy implementation processes that have come to be expected with self-hosted databases. It often takes just a few clicks to set up a cloud-based database in a DBaaS environment, especially with classic self-service use cases in DevOps environments.

- **High agility:**
  Greater agility goes hand in hand with high speed, flexibility and scalability. The rapid provisioning and deployment of DBaaS provides companies with the ideal preparation for dynamic market environments. For example, software developers can set up new databases via web portals or APIs in just one click – instead of having to send tedious enquiries to IT departments.

- **High service quality:**
  The DBaaS operator ensures the optimum performance of the systems according to the SLAs concluded with the company. In addition, companies can conclude SLAs with their managed service provider to ensure compliance and security requirements are consistently observed.

- **Access from anywhere:**
  By moving their databases to the cloud with DBaaS, users can manage every aspect of their database from anywhere with an Internet connection – in the office, at home or on business trips.

6 What are the potential disadvantages of cloudification services?

- **Latencies:**
  When accessing cloud-based databases, it all comes down to the available speed. High latencies can cause accessibility problems, especially for companies that operate in regions with low bandwidths, unless the applications accessing the database are located in the same cloud as the databases themselves.

- **Third-party infrastructure:**
  When opting for DBaaS service providers, companies must be aware that their data will no longer be stored in their own data centres – it will be located in third-party infrastructure. Therefore, they must ensure that the providers will observe all compliance and data protection regulations that apply in their home country, especially when using providers from non-European countries.

- **Data privacy:**
  European companies must pay special attention to the storage and processing of personal data by DBaaS providers to make sure the provisions of the GDPR are consistently observed. This particularly applies to companies in regulated industries (e.g. finance and healthcare). They should exercise particular caution if their data is processed by companies headquartered in the USA, because such companies are subject to the US CLOUD Act and can therefore hardly ensure compliance with the GDPR.

- **Lack of physical control:**
  When using a managed service provider, customers no longer have direct control over the technology used (e.g. security features) or the speed at which new products are introduced. When choosing a provider, it is therefore worth checking how innovative they are, analysing the potential roadmap and concluding appropriate SLAs.

- **Possible vendor lock-in:**
  When choosing a DBaaS provider, companies must ensure that the technologies and systems used are compatible and interoperable with those of other providers; otherwise, they run the risk of getting trapped in a costly vendor lock-in that can quickly negate the benefits of DBaaS. With this in mind, open-source variants should be preferred over proprietary solutions because they can be operated by different providers as open standards, giving companies hugely important room for manoeuvre in the long term.

- **Dependency on the provider:**
  As with most managed services, DBaaS makes companies somewhat dependent on the provider – especially in the areas of performance, security and compliance. That's why companies have to choose the most suitable provider for their purposes – one who meets all their requirements and can react to any challenges that may arise thanks to its fast and excellent service.

- **Pricing and modularity:**
  Some DBaaS providers are increasingly offering "bundles" containing features that users don't need – or rarely need – in their everyday work. That's why companies should make sure they choose providers with the most modular range of products and services, so that they only pay for the services they actually need.

7 What types of database are suitable for the cloud?

There are basically two types of databases, and these differ primarily in the mode of data storage: relational databases (Structured Query Language: SQL) and non-relational databases (NoSQL).

7.1 Relational databases (SQL)

Relational databases are ideal for storing and evaluating structured data. They have been the most common system ever since they were invented in the 1970s. The data contained in relational databases is arranged in tables and can generally be used with Structured Query Language, a programming language for developing and using databases.

The values stored in the table rows are correlated with information from other databases using keys. Companies often use relational databases for transactional processes and web applications (e.g. accounting, controlling, ERP systems, web development and content management).

Some examples of SQL databases include MySQL, PostgreSQL, MS SQL Server, MariaDB, SAP Hana and Oracle Database.

Advantages

- ***High robustness and consistency:*** *As SQL databases have a fixed structure, they are less prone to errors and offer consistently reliable analysis. The tabular arrangement also reduces duplicates.*
- ***Vertical scaling:*** *SQL databases can be scaled vertically in the same system (i.e. by increasing computing power and storage space).*

Disadvantages

- ***Price:*** *SQL databases are often more expensive to run than non-relational databases.*
- ***Difficult horizontal scaling:*** *It isn't too easy to scale SQL databases horizontally (i.e. by connecting additional systems). Additional read replicas are required to increase read access. An active-active configuration set-up is needed to scale write access.*
- ***Limited uses:*** *As SQL databases are only compatible with structured data, they cannot be used in a large number of situations – or only with limited efficiency. For example, they cannot be used to store unstructured data.*

7.2 Non-relational databases (NoSQL)

Non-relational databases are the best choice for organisations that want to store and evaluate unstructured data. An increasing number of companies are using NoSQL databases, because more and more data is available in this form (e.g. documents, images and similar data). As the name suggests, NoSQL databases are not based on the SQL database language, nor do they have a tabular format. Instead, the data stored in NoSQL databases is often organised using objects, documents, lists and rows or value pairs. They are ideally suited to big data applications, because they do away with rigid specifications, schemes and horizontal scaling options.

Some examples of NoSQL databases include MongoDB, Couchbase, Redis, InfluxDB, M3DB, Cassandra and HBase.

Advantages

- ***Flexible data models:*** *In contrast to relational databases, NoSQL databases do not work according to the rigid row-column scheme; they are suitable for a wide range of data structures and can store and process almost all data formats (i.e. structured, semi-structured or unstructured).*
- ***Horizontal scalability:*** *While relational databases only have one server available, additional servers can be added to NoSQL databases to increase their computing power and storage space. This means they are also suitable for big data applications and scenarios in which high peak loads are foreseeable.*
- ***High cost efficiency:*** *As NoSQL databases can be scaled horizontally, no expensive investments have to be made in dedicated central server and storage resources – they can be easily and flexibly booked as required.*
- ***Cloud native:*** *This makes NoSQL databases ideally suited to being run in the cloud, where additional capacities can always be added at the push of a button. That's why the advantages of non-relational databases – in terms of elasticity, flexibility and scalability – are best exploited in a cloud-based environment.*

Disadvantages

- ***Increased manpower:*** *Running NoSQL databases can lead to an increased consumption of resources compared to relational databases, especially in the areas of deployment, configuration management and support – and performance losses caused by lags and other factors can quickly lead to major problems. That's why it is a good idea to find a DBaaS provider who can take care of these tasks and reduce overheads, especially when running non-relational databases.*

7.3 Combination of NoSQL and SQL

It is extremely rare for NoSQL databases to be used on their own. At the very least, companies usually have legacy systems with access to SQL databases. This often includes in-house applications such as ERP systems, HR tools and systems used to control production processes. As SQL databases are also suitable for such purposes, there is often no need to switch to NoSQL. When searching for a managed service provider for the management of databases, companies should therefore make sure that the provider can model both relational and non-relational databases in its environment.

Alternatively, multi-model databases can be produced by putting together different data models in one database. These support different models (e.g. key value, graph, column and relational schemes) and are therefore ideally suited to storing and processing different types of data.

8 What are the applications of DBaaS?

DBaaS can be used in a wide range of scenarios. Some of the most popular use cases are listed below.

Use case	Description	Type of Database	Product
Storage of structured data	The storage and evaluation of data in tabular form (e.g. contact details of customers and suppliers)	Relational	PostgreSQL, MySQL, MS SQL Server, etc.
Storage of unstructured data	The storage and evaluation of unstructured data (e.g. documents, images, key value pairs, diagrams)	Non-relational	MongoDB, Couchbase
Distributed cache service	The use of a database as a cache server (e.g. to accelerate website loading times and carry out real-time analysis, message queuing and high-speed transactions)	Non-relational	Redis
Time series database	The storage and evaluation of repeated measurements (e.g. meteorological data, performance data, medical data, network logs)	Non-relational	MongoDB, etc.
Search engine	An open distributed search engine for the quick retrieval of data (e.g. text data, numerical data, geo-data)	Non-relational	Elasticsearch
Distributed message service	A messaging system based on reliable message queuing	n / a	Apache Kafka
Website activity tracking	The collection, storage and analysis of user activities on a website (e.g. page views, clicks, searches)	n / a	Apache Kafka
Messaging system	The transmission of messages via a message broker (e.g. to separate data processing from data generation and buffer unprocessed messages)	n / a	Apache Kafka, RabbitMQ, ActiveMQ
Metrics / system monitoring	The storage and evaluation of monitoring data from different sources and the aggregation of statistics from distributed applications to create a centralised feed of application data	n / a	Apache Kafka
Logging / log aggregation	The collection and processing of log data from various sources	n / a	Apache Kafka

9 What are the technical challenges?

9.1 Migrating databases to the cloud

When moving databases to the cloud, good preparation is essential. Companies should first check whether the database software of the target cloud is even compatible with their current on-premise software – and whether the size or scope of the database is supported by the provider. Some DBaaS providers only offer smaller database configurations in terms of their storage space and the number of cores used. Companies should clear up these aspects right at the beginning to avoid any nasty surprises.

In addition, a comprehensive assessment of the existing data should be carried out before migration. This will help ensure that the new environment is used as efficiently as possible, particularly with regard to the tariff models applicable in the cloud. As these are determined by CPUs, storage and IOPs, moving unused data can result in unnecessary costs. Therefore, any data sets that are no longer required due to their inferior quality or the expiry of legal deadlines should be sorted out in advance. If older data may still be needed in the future, on the other hand, it can be inexpensively archived in a read-only format, so that it can be used again on demand for business intelligence, machine learning and predictive analysis.

Once this evaluation is done, it is worth thoroughly checking the data to be migrated to the DBaaS environment to eliminate any properties that could impair database performance. This may include uniform naming standards for objects and the absence of oversized columns, MAX-sized data types or implicit conversions.

Existing functions, such as the creation of temporary tables, should also be checked, as they can quickly have a negative impact on performance.

9.2 Important criteria when choosing a DBaaS provider

Once a company has decided to enjoy the benefits of a DBaaS solution, as outlined above, it's time to find the right managed service provider. It is worth considering criteria that are effective in both the short and long term, especially during complex migration projects, in order to ensure the most efficient use of the resources associated with the move to the cloud.

- ***Support for the desired database type or multi-model support***: Companies must first clarify whether the DBaaS provider has the desired system (relational / non-relational) and different models (e.g. key value, graph, column model) in its portfolio. They should make sure that not only proprietary database systems are available, but also especially open-source systems.
- ***Automation for greater efficiency and lower error rate:*** This may include functions such as the automated provisioning of new databases, automatic scaling and diagnostics – and even serverless operations. Such functions increase the efficiency of database operations by significantly reducing the database management workload.

- ***Simple administration and interface***: The relevant interfaces should have a simple and intuitive design to streamline processes. The aim is to minimise the administrative outlay for the operation and monitoring of databases.
- ***Flexibility:*** The chosen DBaaS solution should enable flexible integration in different set-ups and multi-cloud environments – and it should have the necessary APIs for other tools – especially in hybrid environments and when different cloud providers are used. This allows companies to optimise their ability and performance as efficiently as possible while preventing an expensive vendor lock-in.
- ***High availability and resilience:*** Always on should be the standard here. Downtimes have to be reduced to a minimum, especially with DBaaS, as companies otherwise run the risk of losing precious data, which can have serious consequences in the digital economy. A system's resilience is ensured not only through IT security, but also the physical protection of databases in ISO-certified, high-performance data centres and other aspects such as geo-redundancy. This should be recorded in appropriate service-level agreements which, if necessary, can be tailored to the specific customer's needs.
- ***Data security:*** In addition to IT security, data encryption is another important criterion. The encryption of data in transport and in the database should be standard practices for the DBaaS provider, as should the necessary monitoring and key management. Additional data security can be ensured by other services, such as dynamic data masking and the use of tokens.
- ***Ecosystem:*** Cloud-based databases cannot realise their true potential without extra tools and frameworks, such as those required for the integration of data, data pipelining, the implementation of data governance guidelines and other purposes. These tools and frameworks should be provided by the company itself or a strong network of partners to maximise the benefits of DBaaS.
- ***Support:*** If any problems arise, 24/7 technical support should be immediately available – via email or over the phone.
- ***Compliance and data protection:*** When choosing a DBaaS provider, European companies in particular must ensure consistent compliance with legal requirements such as the GDPR – especially with regard to the processing of personal data. Ever since the US CLOUD Act entered into force and the EU-US Privacy Shield was subsequently declared invalid in May 2020, companies must remember that DBaaS providers based in the USA may not ensure the adequate level of protection required by the GDPR.

10 Which providers are there?

A fundamental distinction can be made between the providers on the market: open-source databases provided as a service and proprietary solutions from cloud providers.

10.1 NoSQL (open source)

- **Couchbase:** A document-oriented, non-relational database which stores information in the form of JSON documents.
- **MongoDB Atlas:** A fully managed cloud database service with strong automation functions, high availability and flexible scalability.
- **Redis:** An in-memory database based on the key value principle.
- **InfluxDB:** A database management system designed specifically for recording time series (e.g. for IoT applications or scientific measurements).
- **M3DB:** A distributed time series database which was originally developed by Uber; it can be used for all monitoring applications.
- **Apache Cassandra:** A distributed database management system designed to ensure high scalability and reliability.
- **Apache Hbase:** A key value database for the processing of large volumes of data in the Apache Hadoop ecosystem.

10.2 SQL (open source)

- **PostgreSQL:** An object-relational database management system which can manage alphanumeric data types (letters, numbers and special characters) and complex data types.
- **MySQL:** A relational database management system which was originally developed by Oracle for the management of structured data.
- **MS SQL Server:** A relational database management system developed by Microsoft which is mainly used in data warehouse and business intelligence environments.
- **MariaDB:** A database management system which emerged from a MySQL fork and has its own additional features and extensions.

10.3 SQL (proprietary)

- **SAP Hana:** A technology and development platform based on in-memory technology which is suitable for evaluating large volumes of data.
- **Oracle Database:** A long-established and widely used relational database management system which is particularly used for business applications, data warehousing and business analytics.

10.4 Others (open source)

- **Apache Kafka:** A software solution which was originally developed by LinkedIn as a message queue that enables the storage and management of data streams. Apache Kafka is ideally suited to big data applications, as it has a decentralised architecture, high fault tolerance and scalability.

10.5 Which provider is particularly suitable for ensuring data sovereignty?

In addition to the solutions outlined above, many other database management solutions have been developed by established cloud providers such as Amazon Web Services, Microsoft and Google. They usually come with a large number of technical features and settings. However, the solutions presented here are also ideally suited to the vast majority of use cases involving the storage and evaluation of data.

Proprietary solutions are usually designed in such a way that there are not only technical dependencies between the system components, but also procedural or contractual dependencies. These form a closed system that is mostly incompatible with other solutions on the market. As a result, companies find it incredibly difficult and expensive to migrate their databases to other systems (e.g. to cut costs or ensure compliance). On the other hand, open-source systems offer further significant advantages. In order to avoid such a vendor lock-in, companies should carefully consider which DBaaS provider or database technology will enable them to streamline their processes in the short term – and keep them agile in the long term.

10.6 US CLOUD Act: European providers offer protection

A vendor lock-in isn't the only risk when using DBaaS providers headquartered in the USA. In accordance with the US CLOUD Act, companies based in the USA are obliged to grant the American executive access to the data they host upon request – regardless of the continent in which the relevant data centre is located. The EU-US Privacy Shield, which is used to regulate transatlantic data traffic, was declared invalid by the European Court of Justice in June 2020. Therefore, in accordance with the GDPR, the personal data of EU citizens must not be transmitted to the USA. Any companies that hire US-based DBaaS providers to store and analyse personal data are violating the provisions of the GDPR – and may even risk severe fines[8].

But personal data isn't the only thing companies have to be careful about when storing their valuable data in US clouds.

When companies hire American DBaaS providers, not even their sensitive company secrets, such as those contained in databases used for product development, are safe from the US CLOUD Act. With this in mind, companies should favour European providers who are fully subject to the stringent European data protection provisions

[8] More information can be found in IONOS Cloud's white paper entitled "The controversial CLOUD Act", accessible online at: https://cloud.ionos.co.uk/white-paper/cloud-act.

11 What is the ideal DBaaS product?

The following section outlines some of the key characteristics of high-performance and consistently secure database services.

Access: Users should be able to access the databases hosted in the cloud via private networks and Internet connections, depending on their requests and requirements.

Place of use: This also depends entirely on the needs of the company using DBaaS, especially with regard to its compliance requirements. Both public and private cloud hosting should be possible without any issues. The DBaaS provider should also be able to host the databases in the user's data centre – or enable a high level of performance and configurability via a dedicated bare metal server.

Performance: DBaaS should minimise latencies to ensure that the IOPS value of the storage used remains at a continuously high level. In addition, both horizontal and vertical auto scaling should be integrated in the DBaaS product, so that additional capacities can be accessed when the workload increases – and any resources that are not needed can be switched off to avoid incurring unnecessary costs.

DBaaS providers should offer their users a choice of different storage performance classes to meet the exact performance requirements and maximise cost efficiency. In addition to the HDD and SSD standards, it is also worth enabling SSD premium storage so that databases with high performance requirements can also ensure the required performance.

Availability and reliability: DBaaS providers should have sophisticated failover solutions, so that they can switch seamlessly to a replacement system in the event of failures to minimise downtimes, which can quickly cause users to lose huge volumes of data.

In addition, read replicas should be used with redundant central database instances. These increase both availability and performance: The reading throughput and thus the database performance can be significantly increased by trawling through several database copies at the same time.

Other standard features should include regular physical and logical back-ups of the cloud-based databases, including geo-redundancy for both the data back-up and the failover instance upon request.

Data security: The cloudified databases should be accessed via a modern identity and access management system where different roles and authorisations can be assigned. At the VM level, DBaaS providers should ideally be able to run individual customers in isolation in the respective server instance in accordance with the legal requirements and compliance guidelines.

It should be possible to encrypt data at both the transport level (TLS) and storage level and, if needed, across all levels, including client certificate authentication, etc. Trustworthy DBaaS providers regularly request extensive audits at external test centres to certify their level of protection and prove they meet the compliance level required by the user.

For this purpose, companies should also ensure that their chosen database service fully complies with the GDPR (i.e. they should make sure the provider is a European company which is not subject to the US CLOUD ACT and whose data centres are located in Germany or other European countries).

Handling: When it comes to compiling and configuring the necessary resources, visual user interfaces have proven their worth, as users can simply put together the elements they want via drag and drop. Clear, web-based dashboards are practical for everyday operations, as administrators can quickly extract the relevant metrics. Alerts can also be useful when certain database health and performance values are exceeded, because they ensure a fast response to any issues. Self-service portals have become the standard solution for ordering additional resources, as they make it easier to manage required capacities.

Lock, unlock and read-only modes should be available as additional functionalities for the database cluster, so that no changes are made to the set-up during maintenance work or similar events. Other highly practical functions include start / stop, restart, the database cloning of individual sub-areas and the rapid migration of data between individual clusters or data centres.

In addition, external interfaces are very important to integrate external database analysis tools (incl. plug-in manager) and to quickly and easily move existing databases to a DBaaS environment. Of course, the smooth import and export of data should be ensured.

If users want to get their hands dirty and configure databases themselves, an infrastructure as code component should also be available. This allows infrastructure services to be programmed using machine-readable code, which mainly comes into play in DevOps environments.

Service: As databases are often linked to applications that are critical to a company's success, 24/7 support is mandatory for established DBaaS providers. Updates should be installed quickly and regularly to make sure the performance of the databases

remains stable at all times. A high level of security is also ensured by security patches, which should also be implemented quickly.

Professional SLAs are a standard requirement for good DBaaS products to ensure the corresponding performance level (e.g. scope of services, response time, speed). The failure rates should be no more than 0.1% to continuously ensure the highest level of reliability.

12 Bottom line: Database-as-a-Service – made in Germany

The digital transformation is being driven by continuously growing volumes of data, which have to be stored and evaluated as efficiently as possible to ensure that numerous digital business models can function properly. In this context, many companies are looking to move their databases to the cloud as the next logical step in their development – especially considering the well-documented shortage of IT specialists. DBaaS is the ideal solution where hosting, operations and administration are outsourced to a managed service provider to maximise the benefits of cloud outsourcing.

This offers companies significant advantages over traditional database models, especially in terms of flexibility, scalability, cost efficiency and availability. By cloudifying their databases, companies can reduce their rigid CAPEX in favour of flexible OPEX and outsource activities that employees often find tedious, which then frees up precious IT capacities. Depending on each company's needs, DBaaS providers can supply different types of databases to ensure tailor-made solutions for different challenges concerning the storage and evaluation of data.

When choosing a suitable DBaaS provider, companies should not only pay attention to technical excellence, but also the legal aspects. If personal data is at stake, the GDPR stipulates that it simply cannot be stored in non-European clouds. This particularly applies to DBaaS providers based in the USA, because they are subject to the US

CLOUD Act and therefore cannot ensure the protection of personal data. And there are other cases in which companies should carefully consider whether they want to store their valuable data in such an environment – or whether they would rather have their data stored, managed and analysed by a high-performance and secure provider in Europe.

The first steps on the road to a managed database are the most important

If a company is looking for its own database in the cloud, managed service providers are the ideal choice – especially if the company wants to benefit from a strong partner who can offer advice, infrastructure and operations from a single source. The company should start by analysing its needs with the service provider to work out which database architecture is ideally suited to its use case and which resources are required.

The managed service provider will then take care of providing and installing the necessary systems, procuring and managing the necessary software licences and securely configuring the DBaaS. It will also ensure smooth operations and make sure the database is continuously adapted to meet the company's current requirements – nothing will stand in the way of its digital transformation in the field of data management and analysis.

About IONOS

IONOS is the leading European provider of cloud infrastructure, cloud services and hosting services with over eight million customer contracts. Its product portfolio includes everything companies need to be successful in the cloud, including domains, classic websites, DIY solutions, online marketing tools, full-fledged servers and an IaaS solution. Its products and services are aimed at freelancers, traders and consumers, as well as corporate customers with complex IT requirements.

IONOS Cloud is the European cloud alternative part of IONOS. Our range of products around the Cloud Compute Engine includes an IaaS Compute Engine with its own code stack for virtualisation, managed Kubernetes for container applications, a Private Cloud powered by VMware and S3 Object Storage. We offer established SMEs, major companies, regulated industries, the digital economy and the public sector all the services they need to be successful in and with the cloud.

IONOS was created in 2018 following the merger of 1&1 Internet and ProfitBricks, an IaaS provider based in Berlin. IONOS is part of the listed company “United Internet AG” (ISIN DE0005089031). The brands in the IONOS family include STRATO, Arsys, Fasthosts, home.pl, InterNetX, SEDO, United Domains and World4You.

More information is available at cloud.ionos.co.uk

Imprint

IONOS Ltd.
Discovery House 154 Southgate Street
Gloucester GL1 2EX
United Kingdom

Contact IONOS Cloud

Phone +44 333 336 2984
E-mail product@cloud.ionos.co.uk
Website https://cloud.ionos.co.uk/

Executive Board

Hüseyin Dogan, Dr. Martin Endreß, Claudia Frese, Henning Kettler, Arthur Mai, Britta Schmidt, Achim Weiß

Chairman of the Supervisory Board

Markus Kadelke

Copyright