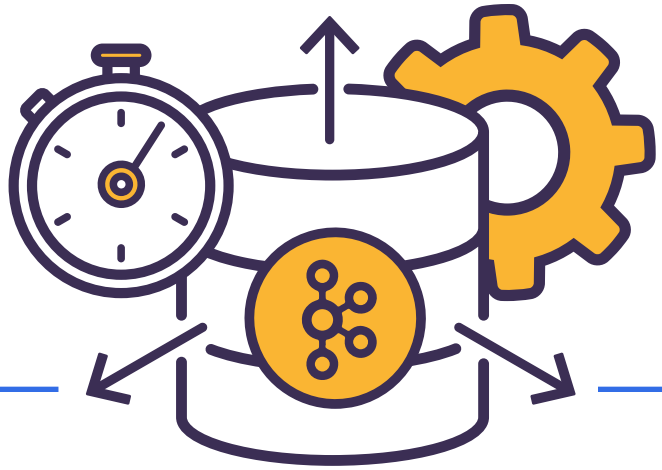
SCYLLA

SCYLLADB WHITE PAPER

7 Essentials When Selecting a NoSQL Database-as-a-Service (DBaaS)



With the move into the “Zettabyte era,” individual data-intensive companies are generating information at the rate of petabytes per day and storing exabytes in total. Some prognosticators believe we’ll see humanity, our computing systems, and our IoT-enabled machinery generating a half a zettabyte of data per day by 2025. All this data presents an unprecedented business opportunity for market disruption.

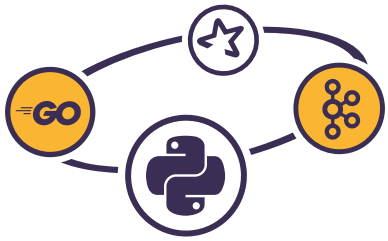
Development teams are enthusiastically building game-changing applications that harness this extraordinary data growth, but few have the time or desire to worry about the database that stores all this data. That’s why so many are looking at database-as-a-service (DBaaS) options: databases that are deployed in a public cloud and fully managed by a DBaaS provider.

A DBaaS can offer advantages such as:

- The ability to get started fast and grow seamlessly
- Simple scaling up and down
- Outsourcing of maintenance/management tasks
- Easy access to top hardware, without having to deal with it
- Guaranteed performance, availability, and security

However, some DBaaS options have drawbacks that can impact long-term agility, productivity, costs, and ultimately growth.

This paper outlines 7 key considerations that help teams realize the many benefits a DBaaS has to offer — without falling into some of the common traps.

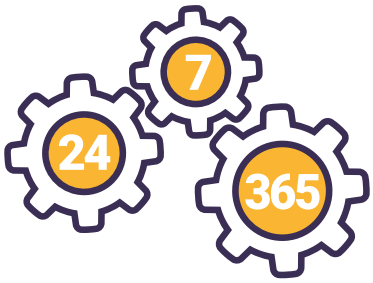


1 A COMPLETE ECOSYSTEM

Look for a DBaaS with a complete ecosystem behind it. This involves support for your preferred programming environment as well as your complete data ecosystem. If you need to build one-off solutions to satisfy your specific use cases, that’s going to drain your resources and inhibit agility as your needs evolve and the ecosystem advances.

First, review all of your application’s data ingress and egress requirements and select a DBaaS solution that supports well-known and established solutions for these data pipelines (Kafka, Spark, Pulsar, Redpanda, etc.). This will result in significant productivity gains and cost savings down the road.

Second, ensure there’s native support for your programming environment. Check if your cloud database service provider publishes a performant, reliable, actively-maintained driver for your specific application languages; this will increase both the client and the DBaaS efficiency.

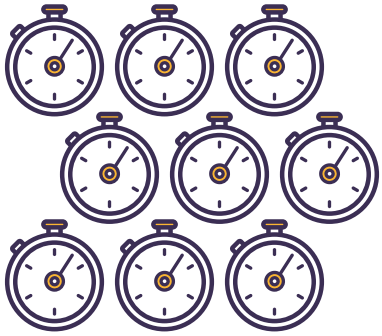


2 ZERO DOWNTIME

Teams often move to a DBaaS solution to guarantee business continuity. Organizations want to ensure zero downtime and rest assured that customers and users can reliably interact with the application 24/7, 365 days a year. To ensure this, look for a DBaaS that's purpose-built for high availability and disaster recovery.

DBaaS solutions that use a peer-to-peer architecture have high resiliency because there's no single point of failure that can bring down the cluster. A peer-to-peer architecture helps even out the workload on the different servers that are deployed underneath the DBaaS solution — without creating bottlenecks or requiring a complex and expensive load balancing system to ensure that the load is well distributed and failover is smooth and fast.

Select a solution that delivers built-in resiliency, a solution that does not force you to purchase premium auxiliary products just for the simple task of having high availability, some type of a connection between the different data centers, and so on. These extra costs can add up fast.



3 REAL-TIME, ALL THE TIME

With users now expecting instantaneous experiences — all the time — the performance of your database is critical. If response speeds matter for your business, look for a DBaaS that can consistently deliver low-single-digit millisecond latencies. Check benchmarks, user experiences, Jepsen test shakedowns, and so forth.

Selecting a DBaaS that provides reliable latency guarantees under the maximum throughput load is especially critical when your business is experiencing record-high transactions and activity. So is making throughput and latency metrics available to your operations team. These metrics will help you perform accurate capacity planning...and plan for future growth as well.

4 COST-EFFECTIVE AT SCALE

No matter where you are today, it's important to anticipate future growth and plan how you can cost-effectively scale your applications to support business growth without causing an exponential increase in costs.

Most DBaaS pricing models are based on operations per second, storage and actual traffic in the cluster. Items such as secondary indexes, if available at all, come at a premium. If you need fast reads, you need to deploy dedicated hardware to support these reads from a specialized cache, which is usually sold at a premium by the DBaaS vendor If your application needs to serve multiple geographical regions, the replication comes at a premium If you scale beyond certain workload, the DBaaS vendor might start to limit your available capacity in certain regions or charge you a premium for having high load on certain regions.

Many DBaaS solutions charge you per write and read units. While these initial charges might seem small in comparison to your storage costs, storage costs are actually the least costly part of the system. For example, assume that one read unit will cost you one-millionth of a cent. When your customer base grows to 10,000 active users per second — each reading one block per second — a year of just those reads will cost you over $3,200 (and that's for only the 10,000 operations per second that occur on that specific cluster). What happens if your user application requires more than one block of reads per second? At that point, just start multiplying those thousands of dollars every time. These read and write operations can become quite expensive over time.

5 AVOID VENDOR LOCKIN

Make sure you're not dependent on a DBaaS solution that prevents you from easily migrating when your needs change or you decide it's time for a different approach. Consider the ability to change cloud providers as well as DBaaS vendors — or even to move from a cloud to on-premises deployment (a process known as cloud repatriation). You might also consider running across multiple public clouds (known as multi-cloud deployments) or span your workload on both public cloud and on-premises deployments (known as hybrid cloud). If these options are important to you, investigate whether you can migrate without changing your application or increasing your total cost of ownership.

Systems like Kubernetes offer the flexibility of selecting infrastructure and moving workloads from one cloud service to another with minimal impact. Selecting a DBaaS that's cloud-agnostic and is easy to use with such systems will make migration operations feasible and save you money down the road.

6 NO UNACCEPTABLE LIMITS

Got large partitions? Wide rows? Huge payloads? Multi-terabyte or petabytes of data? No problem...as long as you find a DBaaS solution that's not limiting you to a narrow scope of workload or payloads.

Different DBaaS services have different read and write characteristics and capabilities. Some even limit the size of your data items or the total storage you're allowed. Can your application work effectively within those boundaries?

Ensure that your DBaaS solution provides operational continuity and scale under the extreme throughput, latency, and storage requirements that you'll see as your application grows and grows.

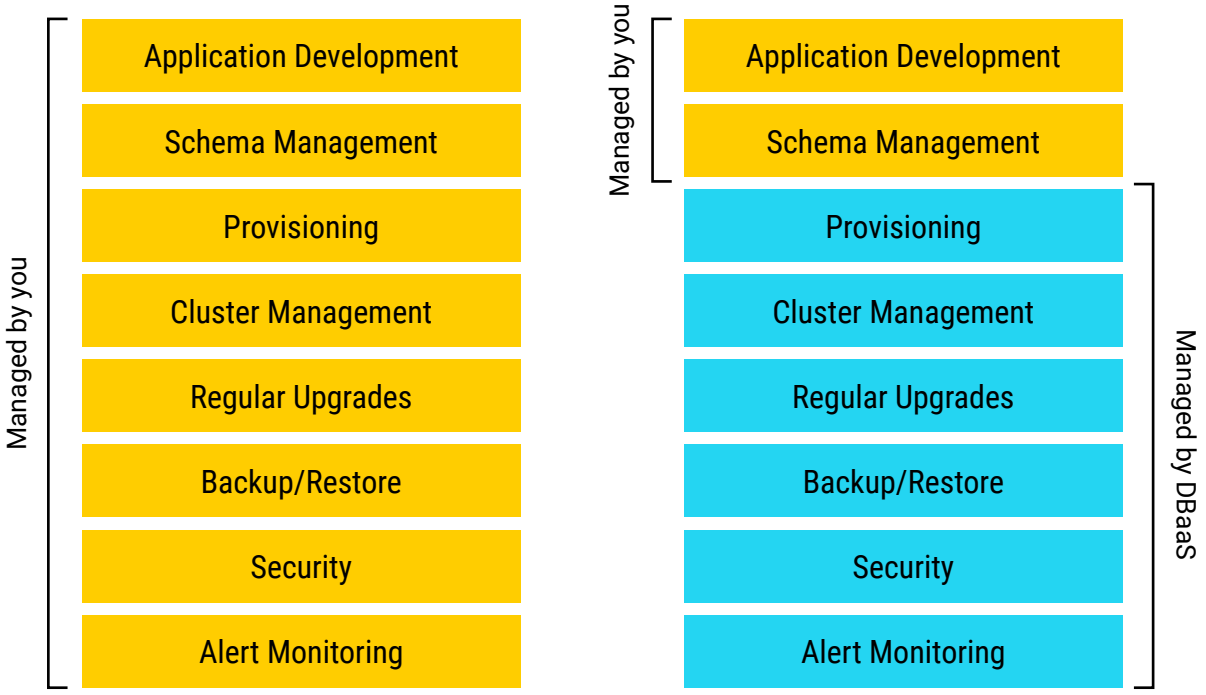
7 OBSERVABILITY

Optimizing your applications requires insight into how the database reacts to your application and whether your application is triggering healthy or unhealthy behavior in the database. Thus, your DBaaS should offer a monitoring capability that not only provides information about your DBaaS, but also reports data about your application's efficiency while using that DBaaS. This helps you discover data patterns that might cause, for example, parasitic latencies, increasing costs, and even allowing malicious access.

Look for capabilities across metrics collection, alerting, dashboards, and log aggregation. Simple integration with popular tools like Grafana, Prometheus, and Datadog is also helpful. All-around effective monitoring and observability functionality will help DevOps, infrastructure operations teams, and database administrators quickly find and fix issues impacting the performance.

ADDITIONAL CONSIDERATIONS

One of the greatest advantages of a DBaaS is that it relieves you of the time and hassle of dealing with your database. That way, you can focus on your application. Here's a look at what you should plan to continue managing, and what you can expect your DBaaS vendor to handle.



What you are responsible for with a cloud-hosted self-managed database (left) versus with a database-as-a-service (right).

Also, be sure that you have sufficient internal resources to support the DBaaS evaluation and adoption process. Here is a general guide on what you should plan for:

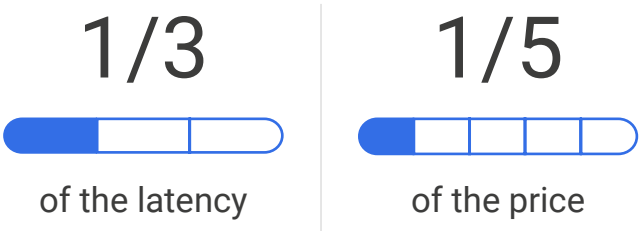
Item	Expected Duration	Involved Stakeholders	Target	Succes Criteria
Use case review	2-3 weeks	Business, Application, Infrastructure, Evaluators	Proof of Concept document	Buy in from all stakeholders on PoC document
System Installation	1-2 weeks	Infrastructure, Evaluators	Working Setup, well configured loaders and monitoring	Database installation completion
Resilience tests	1 week	Infrastructure, Evaluators	Create disaster events measure cluster availability	Application layer stay uninterrupted during stress
Re-tune systems	1 week	Application, Infrastructure, Evaluators	Improve setting based on benchmark learnings	Optimized deployment

INTRODUCING SCYLLA CLOUD

One option for satisfying these essentials is to take advantage of Scylla Cloud, a DBaaS featuring ScyllaDB.

ScyllaDB is the database for data-intensive apps that require high performance and low latency. We enable teams to harness the ever-increasing computing power of modern infrastructures — eliminating barriers to scale as data grows. Unlike other databases, ScyllaDB is built on extreme architectural advancements so you can thrive in this next tech cycle.

Scylla Cloud is a fully managed version of ScyllaDB allowing you to focus on your applications, sparing your team from performing all database administrative tasks such as automatic backups, repairs, monitoring, performance optimization, security hardening, plus 24*7 maintenance and support — all for a fraction of the cost of other DBaaS solutions.



Scylla Cloud delivers consistent, low-single-digit latencies. A recent Yahoo! Cloud Serving Benchmark (YCSB) shows ScyllaDB p95 latency as one-third that of Amazon DynamoDB. Our prices are even lower — as little as one-fifth of what you'd pay for the same workload on DynamoDB.

Here's a quick look at why game-changers like Disney+ Hotstar, Medium, and Crypto.com rely on Scylla Cloud to power their data-intensive applications:

- **Price Performance:** We achieve millions of OPS throughput from a single node, resulting in huge cost savings.
- **Low Latencies:** Consistent single-digit millisecond P99 latencies ensure dependable performance.
- **No Vendor Lock-In:** Built for portability, you can easily migrate data across different clouds or on-premise deployments.
- **Cassandra / DynamoDB Compatible:** Migrate to Scylla Cloud from Apache Cassandra or Amazon DynamoDB without changing your application code.
- **Managed by the Experts:** ScyllaDB engineers who developed the software continuously monitor system health and provide 24*7 online support.
- **Flexible Pricing & Sizing:** On-demand or yearly commitments, with the added option to run on your own cloud account.
- **Scales Up and Out:** ScyllaDB's performance grows linearly with larger compute instances and additional cores.
- **Just a Better NoSQL Database:** Handles hot partition keys 20X better than DynamoDB/BigTable/Cosmos DB. Each partition can be up to 100GB in size, relieving you from changing your data model. Equal write and read performance, all internal nodes are active-active. Smart clients reach the target shard in a single hop.

ABOUT SCYLLADB

ScyllaDB is the database for data-intensive apps that require high performance and low latency. It enables teams to harness the ever-increasing computing power of modern infrastructures – eliminating barriers to scale as data grows. Unlike any other database, ScyllaDB is built with deep architectural advancements that enable exceptional end-user experiences at radically lower costs. Over 400 game-changing companies like Disney+ Hotstar, Expedia, FireEye, Discord, Crypto.com, Zillow, Starbucks, Comcast, and Samsung use ScyllaDB for their toughest database challenges. ScyllaDB is available as free open source software, a fully-supported enterprise product, and a fully managed service on multiple cloud providers. For more information: ScyllaDB.com

SCYLLADB.COM

SCYLLA

United States Headquarters
2445 Faber Place, Suite 200
Palo Alto, CA 94303 U.S.A.
Email: info@scylladb.com

Israel Headquarters
11 Galgalei Haplada
Herzelia, Israel