



Scylla Cloud

The powerful Scylla NoSQL database
as a highly affordable cloud service



CONTENTS

REAL-TIME BIG DATA IN THE CLOUD	3
INTRODUCING SCYLLA CLOUD	3
WHY SCYLLA CLOUD?	4
PREDICTABLE LOW LATENCY	4
PRICE TO PERFORMANCE	5
NO VENDOR LOCK-IN	5
MANAGED BY PROS	5
SCYLLA CLOUD FEATURES	5
VERTICAL AND HORIZONTAL SCALE	5
HIGH AVAILABILITY AND OPTIMIZED, CROSS-DATACENTER REPLICATION	6
FLEXIBLE, RESILIENT DATA MODELING	6
HANDS-FREE MAINTENANCE AND UPGRADES	6
HOT KEY PROTECTION	6
CASSANDRA COMPATIBILITY	6
ENTERPRISE-GRADE SECURITY	6
SOME SCYLLA CLOUD USE CASES	7
ADTECH	7
INTERNET OF THINGS (IOT)	7
CUSTOMER 360/PERSONALIZATION	7
PRODUCT CATALOGS/SHOPPING EXPERIENCE	7
FRAUD DETECTION	7
COMPARISONS	7
GETTING STARTED WITH SCYLLA CLOUD	8
SUMMARY	10
ADDITIONAL RESOURCES	10
REFERENCES	10



REAL-TIME BIG DATA IN THE CLOUD

Companies are moving infrastructure to public clouds to reduce operational overhead and increase agility. Recent studies show that almost 90 percent of companies plan to increase budgets for cloud in the next year. Research firm IDC [predicts](#) that by 2025, 49 percent of data will be stored in public cloud environments, and nearly 30 percent of the data generated will be consumed in real time.

Alongside this general trend towards the cloud, NoSQL databases have become the standard data platform for applications that rely on unstructured data. Unstructured datasets are growing quickly and applications that consume it are increasingly popular. Unstructured data feeds many AI and machine-learning algorithms, and is increasingly used for real-time operational applications. By some estimates, unstructured data represents 80% of all data in companies and is growing up to 50 times as fast as structured data. According to [a 2019 survey of IT professionals](#), the typical organization reports unstructured data growing around 25% annually, which means it will double every 40 months.

With this extraordinary growth in both the quantity and value of unstructured data, many organizations are looking for a suitable cloud-based NoSQL solution for their real-time big data needs. To date, a key obstacle to cloud NoSQL adoption has been the fear of valuable data being held captive by an external entity. All other concerns flow from this basic problem.

By now, most IT organizations are aware of the benefits the cloud provides, among them enhanced agility, organizational focus, and the conversion of capital expenditure (CAPEX) to operational expenditure (OPEX). Since organizations no longer need to procure servers, networking and storage, CAPEX is reduced or eliminated. Although funding is still required to rent servers and associated services from a cloud provider, these outlays can be more flexibly calibrated against business requirements.

Yet with datasets growing not only in size but also in value, companies need total control over that data, including the ability to seamlessly move their data back to on-premises datacenters or to another cloud platform as required. They also need to not have to worry that performance issues with a managed service would interfere with the operation of their real-time business.

Performance, defined by low latency and high throughput, is critical for cloud databases, which often fail to deliver on the performance promises made by providers. Many databases can claim to be ‘fast’; but the bigger challenge is to provide high performance that is also predictable. A common measure of this performance is 99th percentile latencies.

Another common goal for many organizations is to maximize their “infrastructure optionality” by minimizing or eliminating vendor lock-in. Many database vendors (especially cloud platforms) restrict their customers to single deployment topology, making them vulnerable to arbitrary decisions from cloud vendors, such as sudden changes in support, pricing, and capabilities.

Cloud databases make provisioning of database infrastructure much easier than on-premises. With a few clicks, operators can add capacity that’s immediately available to connect to existing applications. However, many cloud databases require their users to over-provision capacity, adding expensive resources that effectively remain idle for long periods of time. Cloud databases carry premium pricing, and such over-provisioning only adds to the end-of-month sticker shock experienced by many teams.

INTRODUCING SCYLLA CLOUD

Scylla Cloud delivers the fastest and most reliable NoSQL database as a managed service. It delivers consistent low latency, high availability, and prioritized multi-workload capabilities, all protected by enterprise-grade single-tenant security. Scylla’s unique

architecture takes maximum advantage of the latest hardware, including NVMe, which fully exploits the levels of parallelism available in modern SSDs. This makes Scylla not only faster than other offerings, but also less expensive to run overall, both in terms of hardware provisioned as well as reduced operational overhead.

That said, users of Scylla Cloud can remain blissfully unaware of setup, configuration, and maintenance. By subscribing to Scylla Cloud, users are relieved of the provisioning headaches of on-premises hardware acquisition and the on-going maintenance and upgrades of database software. The managed service automates administrative drudgery such as provisioning, configuration, upgrades, and backups, freeing up teams to be as productive and as agile as possible and to focus on developing modern apps that build the business.

Moving to the cloud is about more than simply reducing costs and overhead. Organizations still need to maintain Service Level Agreements (SLAs) through predictable low latency and high availability. In many cases, teams are looking for a

managed database that can deliver even higher speed and support bigger workloads, while overcoming technical limitations of existing cloud databases. Scylla Cloud meets both of these needs.

WHY SCYLLA CLOUD?

Scylla Cloud runs the latest version of Scylla Enterprise. Like Scylla Enterprise, Scylla Cloud scales up as well as out, enabling users to optimize cluster configuration to meet specific requirements for performance and cost. Scylla Enterprise has been benchmarked at millions of OPS per server, and it scales as more resources are added to instances, and as more servers are added to a cluster.

PREDICTABLE LOW LATENCY

Consistent single-digit millisecond 99th-percentile latencies ensure dependable performance on Scylla Cloud. Scylla Cloud is implemented in C++, enabling the software to take advantage of modern hardware and to

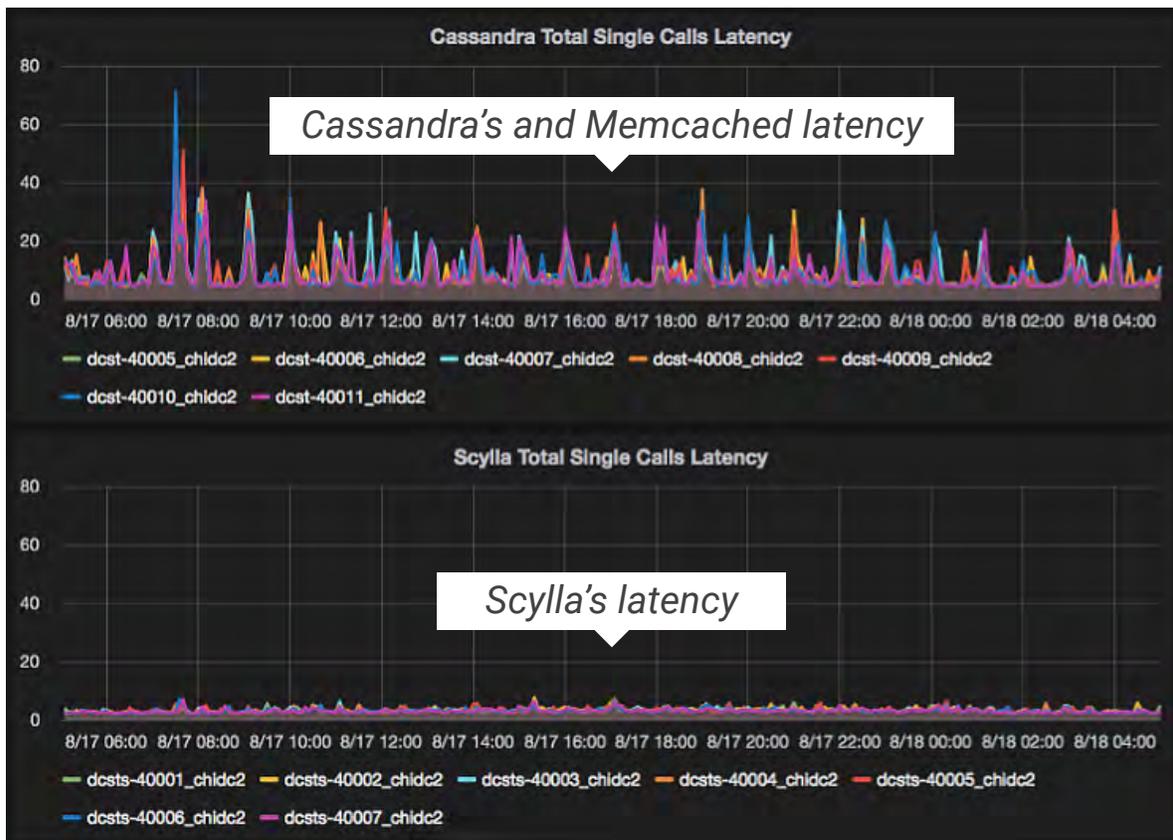


Figure 1 - Scylla's latency compared to Cassandra's and Memcached's latency.

have fine-grained control over the CPU, I/O and memory data transfers. Unlike databases that are written in Java, Scylla Cloud never suspends application processing for garbage collection, compaction or other housekeeping activities that must be performed in virtualized environments. Figure 1 shows how much lower and consistent the Scylla's latency is compared with Cassandra's latency. The Cassandra environment included a caching layer as well.

PRICE TO PERFORMANCE

Scylla's millions of operations per node throughput reduces the number of nodes required to support applications, making Scylla Cloud significantly less expensive than comparable cloud database offerings. Its unique architecture enables it to support larger datasets and heavier workloads with smaller clusters of more powerful servers. This hardware-level efficiency lowers infrastructure costs overall. The result is significant, quantifiable cost savings versus other cloud databases. [See the whitepaper here.](#)

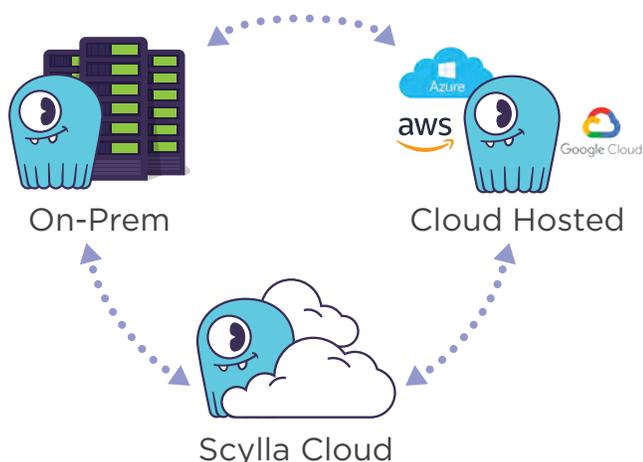


Figure 2 - Deployment options.

NO VENDOR LOCK-IN

Scylla Cloud is built for portability. Organizations can easily migrate datasets across cloud platforms or to on-premise deployments of Scylla Enterprise. Interoperability is supported through a standard and well-documented data format, which enables users to move data between on-premises databases, whether Scylla Enterprise or Apache Cassandra, and public cloud environments such

as AWS, Azure, and GCP. To accommodate data residency requirements, you can store your data in various geographical locations and migrate data between cloud providers as your regulatory and compliance needs demand.

MANAGED BY PROS

To provide Scylla Enterprise in the cloud, ScyllaDB engineers automated the installation, configuration and monitoring of Scylla Enterprise. They manage it on your behalf, seamlessly installing updates, bug fixes, and upgrades, all while your applications continue to run, with zero downtime or service interruptions.

SCYLLA CLOUD FEATURES

VERTICAL AND HORIZONTAL SCALE

Scylla's performance grows linearly with available computing resources. Scylla Cloud supports a range of AWS instances types, and therefore can handle a wide range of application needs. AWS instances offer from 2 to 64 vCPUS, with solid-state drives that provide from .475TB of storage to a 15.2TB of data. Scylla scales easily, meaning that users can start with small instances for development and testing, while migrating to more powerful clusters for testing at scale and even more powerful instances for production deployment. Table 1 details the specific instances.

Instance Name	vCPUs	RAM (GB)	Storage (TB)
t2.micro	1	.5	.005 EBS
i3.large	2	15.25	0.475 SSD
i3.xlarge	4	30.25	0.95 SSD
i3.2xlarge	8	61	1.9 SSD
i3.4xlarge	16	122	3.8 SSD
i3.8xlarge	32	244	7.6 SSD
i3.16xlarge	64	488	15.2 SSD

Table 1 - Specifics of Scylla Cloud instances.

HIGH AVAILABILITY AND OPTIMIZED, CROSS-DATACENTER REPLICATION

Scylla Cloud is available on any combination of AWS regions, delivering on the promise of globally distributed, highly available NoSQL data.

Taking advantage of AWS's globally distributed architecture, Scylla can frictionlessly replicate data across availability zones within a region to eliminate single points of failure while shielding against regional datacenter outages. If an entire zone goes down, other copies of data remain available, and applications will continue to run uninterrupted. Once a zone is back online, Scylla Cloud brings the replica up to date.

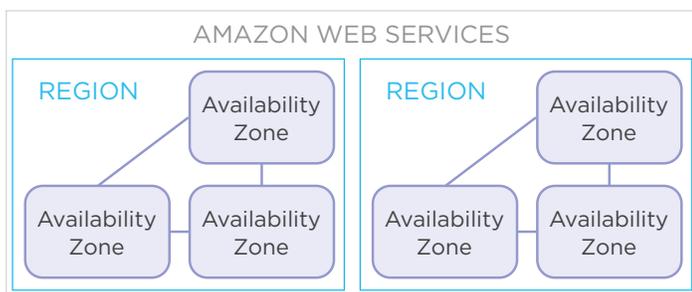


Figure 3 - AWS Regions and Availability Zones.

FLEXIBLE, RESILIENT DATA MODELING

As a wide-column store, Scylla imposes minimal limitations on data models and supports binary large objects (BLOBs). Resilience in the face of hot partitions provides greater flexibility in specifying keys.

HANDS-FREE MAINTENANCE AND UPGRADES

Scylla Cloud always runs the most recent version Scylla Enterprise, that include the latest features and security updates, which are seamlessly applied and allow for upgrades that won't interrupt running clusters. ScyllaDB will install these updates and will always install security patches. In addition, ScyllaDB works with AWS to install system patches as required by Amazon Web Services. This gives you the confidence that you are running the most performant and feature-rich version of Scylla Enterprise in a trusted and safe environment.

HOT KEY PROTECTION

Hot keys, or hot partitions, are a problem for some cloud NoSQL databases, notably Amazon DynamoDB, Google BigTable and Microsoft Cosmos DB. Hot keys can cause traffic to be interrupted by throttling exceptions. In Amazon DynamoDB, for example, global reservation is divided into partitions, each of which is limited to 10 gigabytes. Following the law of "Zipfian distribution," some partitions attract exponentially more queries than others. Such a hot partition can reach throttling limits even when overall traffic is within the global reservation. For this reason, it's necessary to over-provision capacity as a whole to ensure that hot partitions don't impact availability.

Scylla Cloud has reduced the limitation of hot partitions compared to other NoSQL databases. Scylla Cloud handles hot keys 20X better than other cloud NoSQL databases. Scylla partitions can be much, much higher, providing much greater flexibility in data modeling.

CASSANDRA COMPATIBILITY

Scylla Enterprise is a drop-in replacement for Apache Cassandra. The API for Scylla Cloud is the same as Apache Cassandra, meaning that application developers can seamlessly take advantage of the Cassandra ecosystem to maintain productivity while migrating or implementing applications. Examples of this include Kafka, Spark, and Presto, among others. In addition to being able to use existing drivers, Scylla offers optimized drivers that allow applications implemented in many common programming environments to take full advantage of Scylla Cloud.

ENTERPRISE-GRADE SECURITY

Scylla Cloud provides all of the security features of Scylla Enterprise, including role-based access control, encrypted backups, and key management, supported by single-tenant security.

SOME SCYLLA CLOUD USE CASES

Scylla Cloud excels across a range of industries and use cases that call for extremely fast data storage and retrieval and predictable, low latencies. Following are a few use cases that illustrate the advantages Scylla provides.

ADTECH

AdTech thrives on speed. Digital real estate is claimed in milliseconds, matching bids to user profiles. Being able to leverage user attributes such as interests or buying habits requires single-digit millisecond responses from the database. Scylla Cloud is proven to consistently meet the stringent SLAs imposed by real-time bidding platforms.

INTERNET OF THINGS (IOT)

Machine data is growing at an exponential rate. A simple thermostat might generate 50MB of time series data per month. A complex, instrumented turbine blade might generate hundreds of gigabytes of data in a single day. Still, both sensors require ultra-low latencies and high throughput, and large direct attached storage means that data can be ingested at high speed and retrieved for operational usage, monitoring UIs, and analytics applications.

CUSTOMER 360/PERSONALIZATION

A comprehensive customer profile encompassing buying history, social media engagement, CRM interaction, and other relevant data can enable retailers to optimize offers and create more satisfied customers. Scylla Cloud gives organizations the ability to quickly retrieve a range of unstructured data to maximize customer spend.

PRODUCT CATALOGS/SHOPPING EXPERIENCE

Product catalogs change constantly and require information from many sources. With Scylla Cloud, modifying a product catalog by gathering information from a variety of sources and presenting to the customer at interactive

speeds leads to higher customer satisfaction. As shopping patterns vary by season or popular events, Scylla Cloud customers can easily expand or contract their database footprint to meet customer demand.

FRAUD DETECTION

With more financial transactions taking place over the internet, fraud detection algorithms are constantly being developed to counter the threat. To stop fraud in real-time, applications rely on databases to return information in the single-digit milliseconds. Scylla Cloud helps companies meet SLAs when staying ahead of new and sophisticated methods of fraud.

COMPARISONS

Scylla Cloud delivers better performance at a lower price compared to cloud managed databases that are based on non-native programming languages. Scylla Cloud's instance provisioning model provides complete control and visibility over your deployment topology. Unlike other DBaaS systems, where you specify the anticipated throughput limits and then must adjust accordingly at additional expense, Scylla Cloud allows you to understand and manage costs before provisioning the cluster, with no fear of end-of-the-month sticker shock.

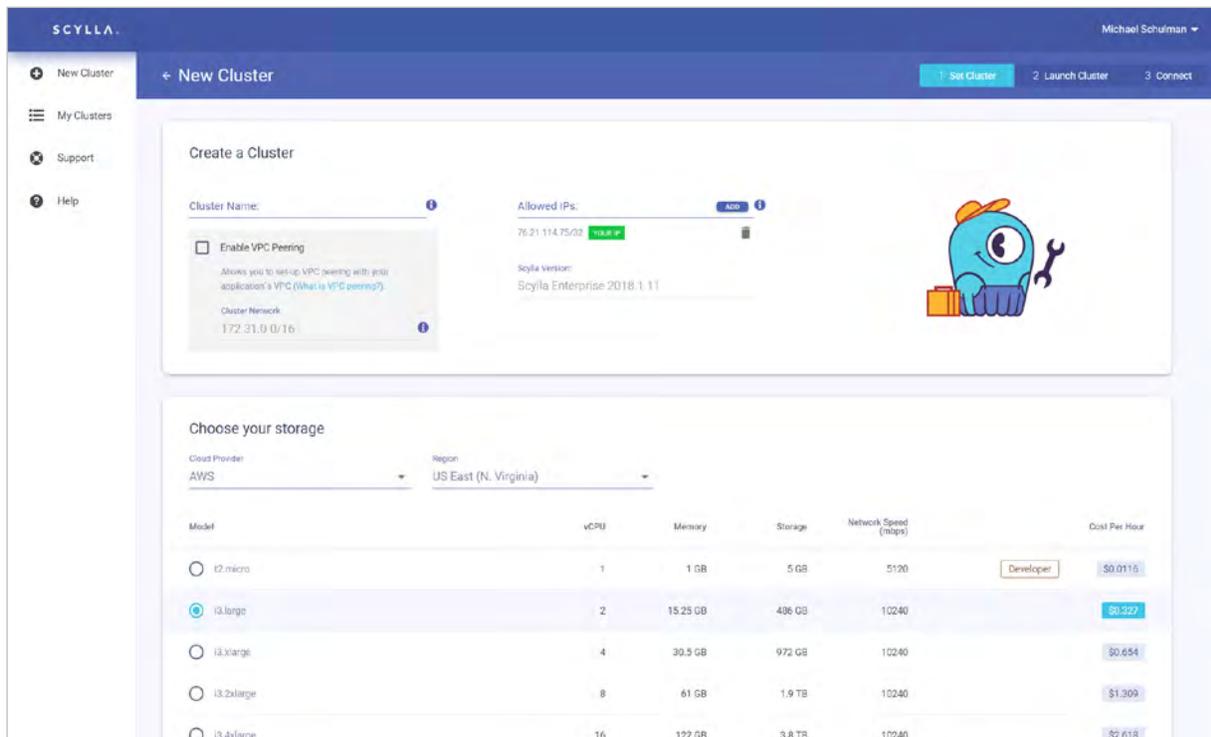
Research shows that Scylla Cloud is one-fifth the cost of the equivalent Amazon DynamoDB throughput. Scylla Cloud also demonstrated 3x-4x better tail latencies than Amazon's DynamoDB. Amazon DynamoDB had issues loading large amounts of data, which resulted in users having to purchase more throughput than was initially specified (an additional cost).

You can find the benchmark at: scylladb.com/2018/12/13/scylla-vs-amazon-dynamodb.

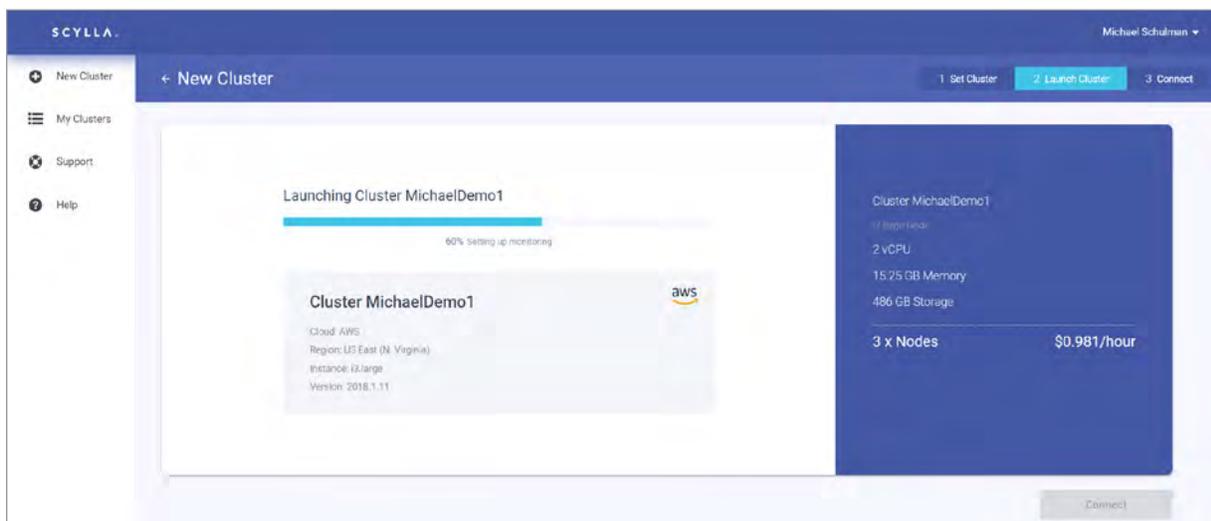
GETTING STARTED WITH SCYLLA CLOUD

Scylla Cloud is quick and easy to provision.

1. Create a Scylla Cloud account at: cloud.scylladb.com/user/signup
2. Log in to your account at: cloud.scylladb.com/user/signin
3. Identify your cluster name as well as the AWS Region where you want your instances to be located and provisioned. Select the type of instance from the list presented. If you wish, you can change your Replication Factor (default = 3) and the number of nodes you need.



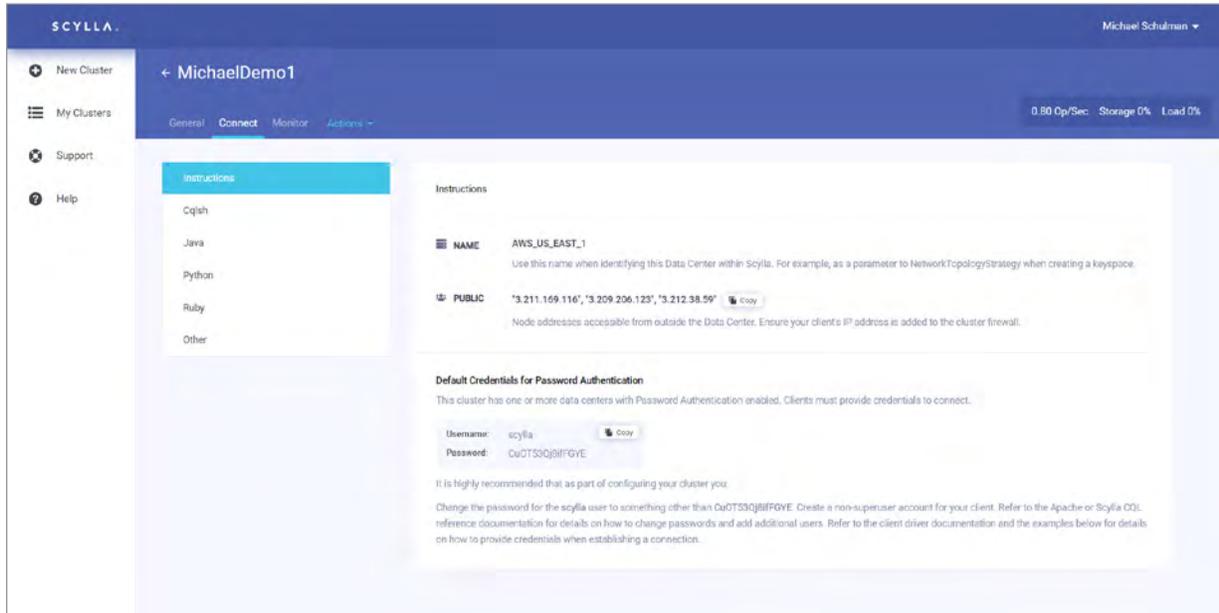
4. Scylla Cloud will dynamically provision the type and number of nodes you specified. This takes just a few minutes, and you will see a screen like the image below, which displays information about your cluster, as it is being provisioned.



5. To retrieve the initial password and the IP addresses for your Scylla Cloud cluster, select **Connect**.

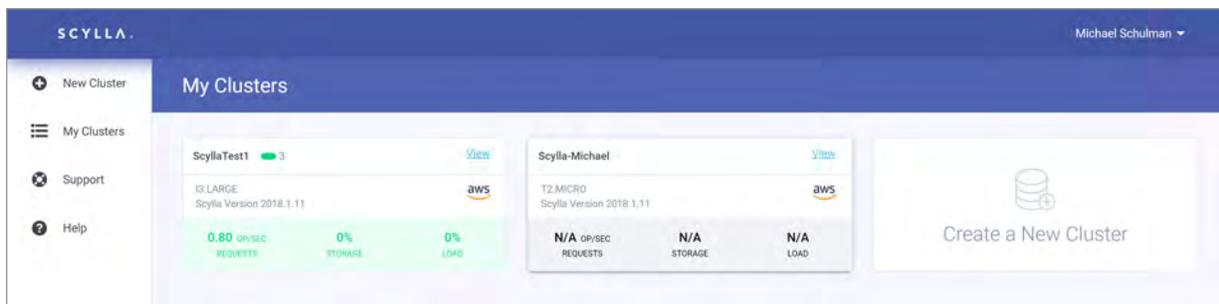
Examples are provided to show you how to connect applications in the following programming languages to your Scylla Cloud cluster:

- Cqlsh
- Java
- Go
- Python
- Ruby
- C#
- C++
- Node.js
- PHP



6. You can view all of your provisioned clusters by selecting **“My Clusters”** on the left-hand menu.

An example is shown below.



SUMMARY

Scylla Cloud delivers a cost-effective, low-touch deployment option for companies looking to build applications using Scylla's real-time big data NoSQL database. It's a flexible and fully managed database service that enables organizations to choose the AWS instances best suited to specific workloads and business requirements. Since Scylla Cloud is managed by Scylla engineers, Scylla Cloud frees up IT organizations to concentrate on developing modern applications to serve their customers, rather than deal with administrative tasks. Scylla Cloud outperforms the most popular DBaaS at a fraction of the cost.

Try Scylla Cloud for yourself. Our Developer Plan enables you to get started for pennies a day. Visit our [Scylla Cloud site](#).

ADDITIONAL RESOURCES

Scylla Cloud web page

scylladb.com/product/scylla-cloud

Scylla Cloud FAQ

scylladb.com/product/scylla-cloud/#faq

Scylla Cloud Pricing

scylladb.com/product/scylla-cloud/#pricing

Scylla Enterprise web page

scylladb.com/product/scylla-enterprise

Scylla Documentation

docs.scylladb.com

Scylla Cloud Blog

scylladb.com/2019/04/09/introducing-scylla-cloud

REFERENCES

- (1) datometry.com/resources/cloud-express-articles/database-migration-challenges-and-truth-about-replatforming
- (2) commvault.com/news/2017/august/study-confirms-c-level-and-other-it-leaders-have-fear-of-missing-out-on-cloud-advancements

ABOUT SCYLLADB

Scylla is the real-time big data database. A drop-in alternative to Apache Cassandra and Amazon DynamoDB, Scylla embraces a shared-nothing approach that increases throughput and storage capacity as much as 10X that of Cassandra. AdGear, AppNexus, Comcast, Fanatics, FireEye, Grab, IBM Compose, MediaMath, Ola Cabs, Samsung, Starbucks and [many more leading companies](#) have adopted Scylla to realize order-of-magnitude performance improvements and reduce hardware costs. Scylla is available in Open Source, Enterprise and fully managed Cloud editions. ScyllaDB was founded by the team responsible for the KVM hypervisor and is backed by Bessemer Venture Partners, Eight Roads Ventures, Innovation Endeavors, Magma Venture Partners, Qualcomm Ventures, Samsung Ventures, TLV Partners, Western Digital Capital and Wing Venture Capital.

For more information: [ScyllaDB.com](https://scylladb.com)

SCYLLADB.COM



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

Israel Headquarters
11 Galgalei Haplada
Herzeliya, Israel

