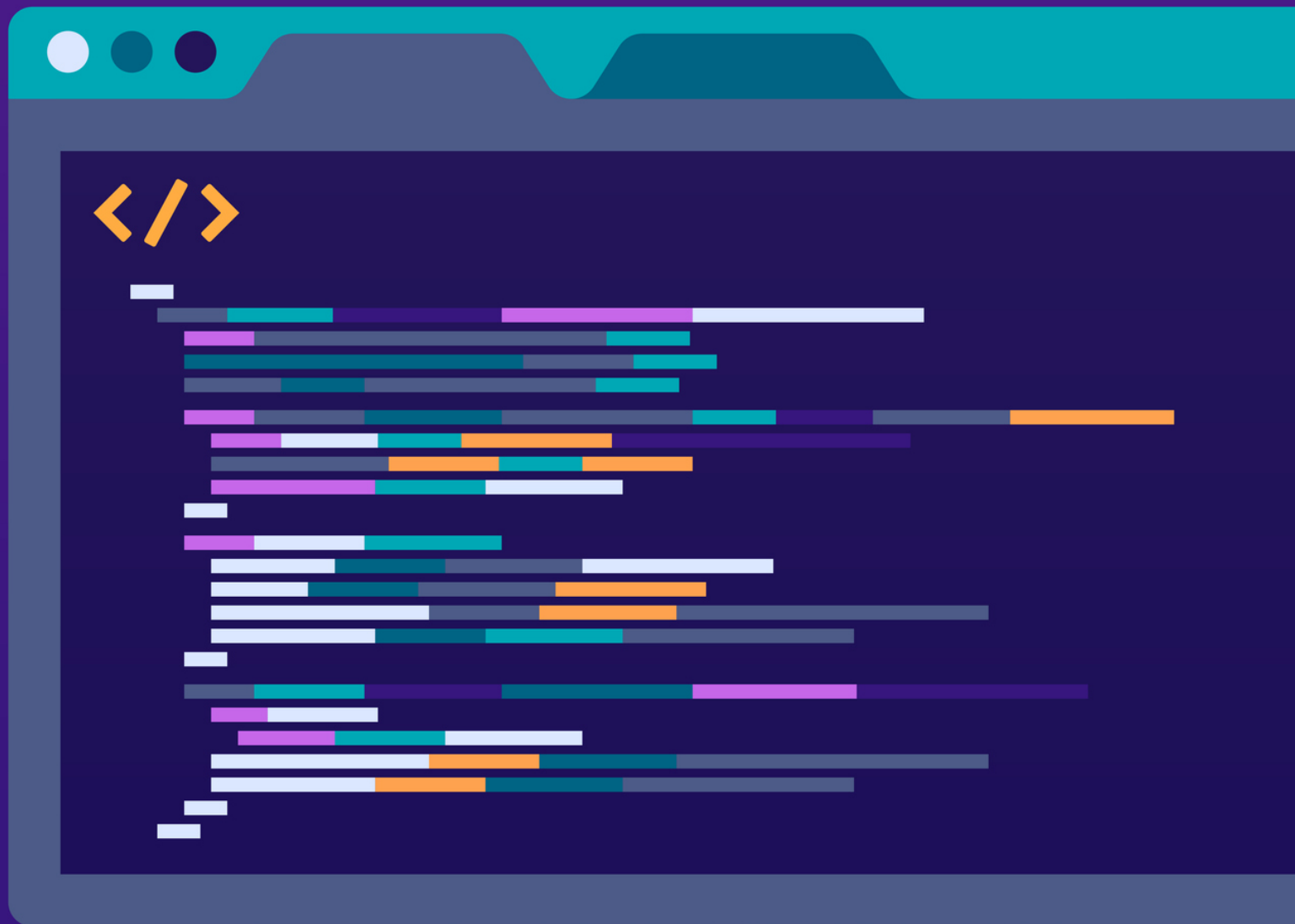



# When, Where and Why to Use NoSQL



A stylized illustration of a city skyline in shades of gray, featuring various skyscrapers and buildings of different heights and shapes. The background is a light teal color with white, rounded, cloud-like shapes.

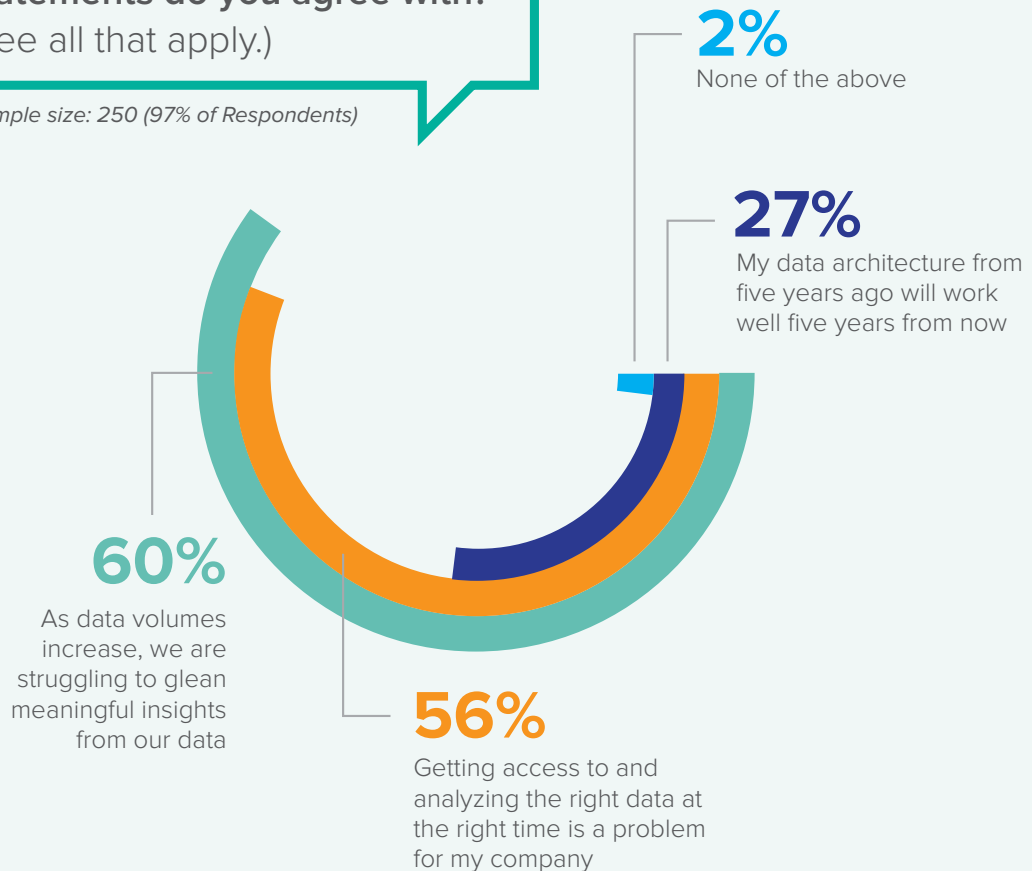
Big data is becoming a big challenge for enterprises. Many organizations have built environments for transactional data with Relational Database Management Systems (RDBMS), but are now inundated with enormous amounts of new types of data. These include, but are not limited to, social media data, server logs, clickstream data, machine data, and geolocation data.

These new data sources include unstructured and semi-structured data, and they all share the common big data characteristics of volume, velocity, and variety.

**Big Data, Big Challenges.** Modern enterprises collect more data than ever before, from a wide variety of sources, and in a wide variety of formats. In a recent survey conducted by Researchscape, executives surveyed reported that as data volumes increase, they are struggling to glean meaningful insights from their data (60%), and that getting access to and analyzing the right data at the right time is a problem (56%). Only a quarter of executives believed that their data architecture from five years ago will work well five years from now (27%).

Which, if any, of the following statements do you agree with?  
(See all that apply.)

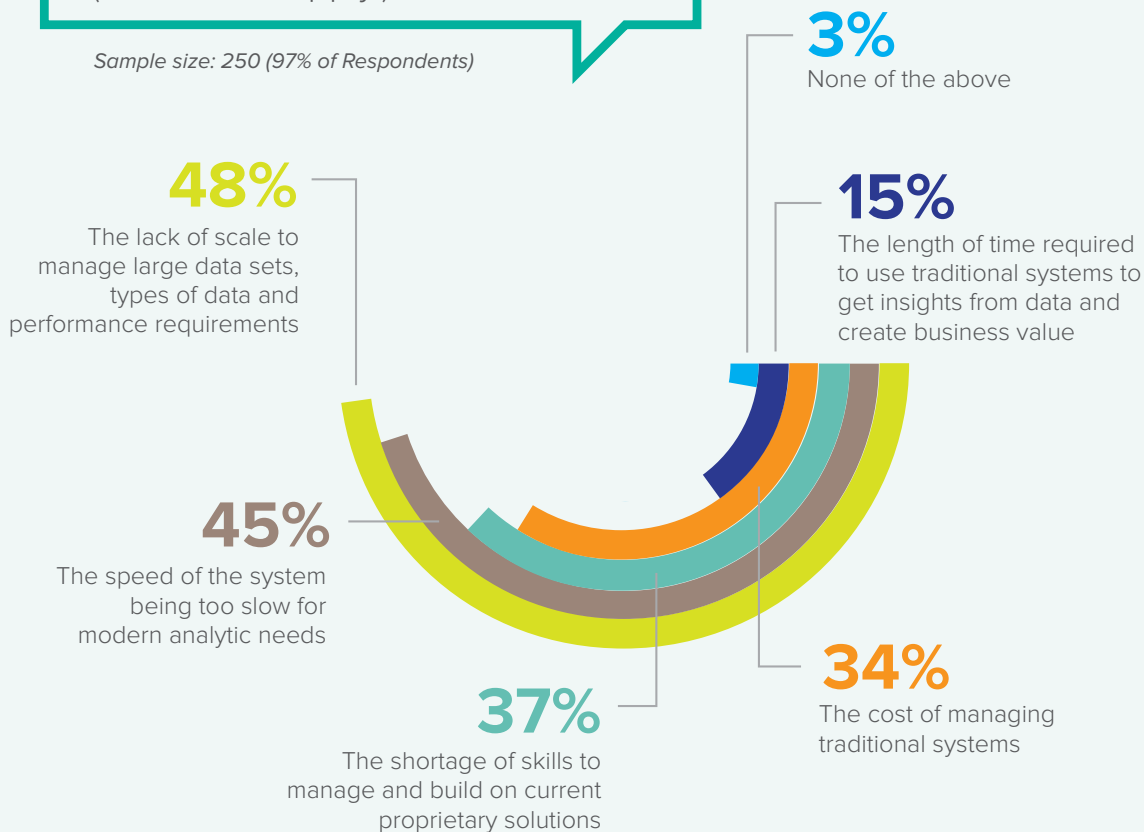
Sample size: 250 (97% of Respondents)



Challenges such as the lack of scale to manage large data sets (48%) and the slow speed of the system for modern analytics (45%) concern the executives surveyed and may drive them to change their traditional data management systems.

Which of the following statements concern you and may drive you to change your traditional data management systems?  
(See all that apply.)

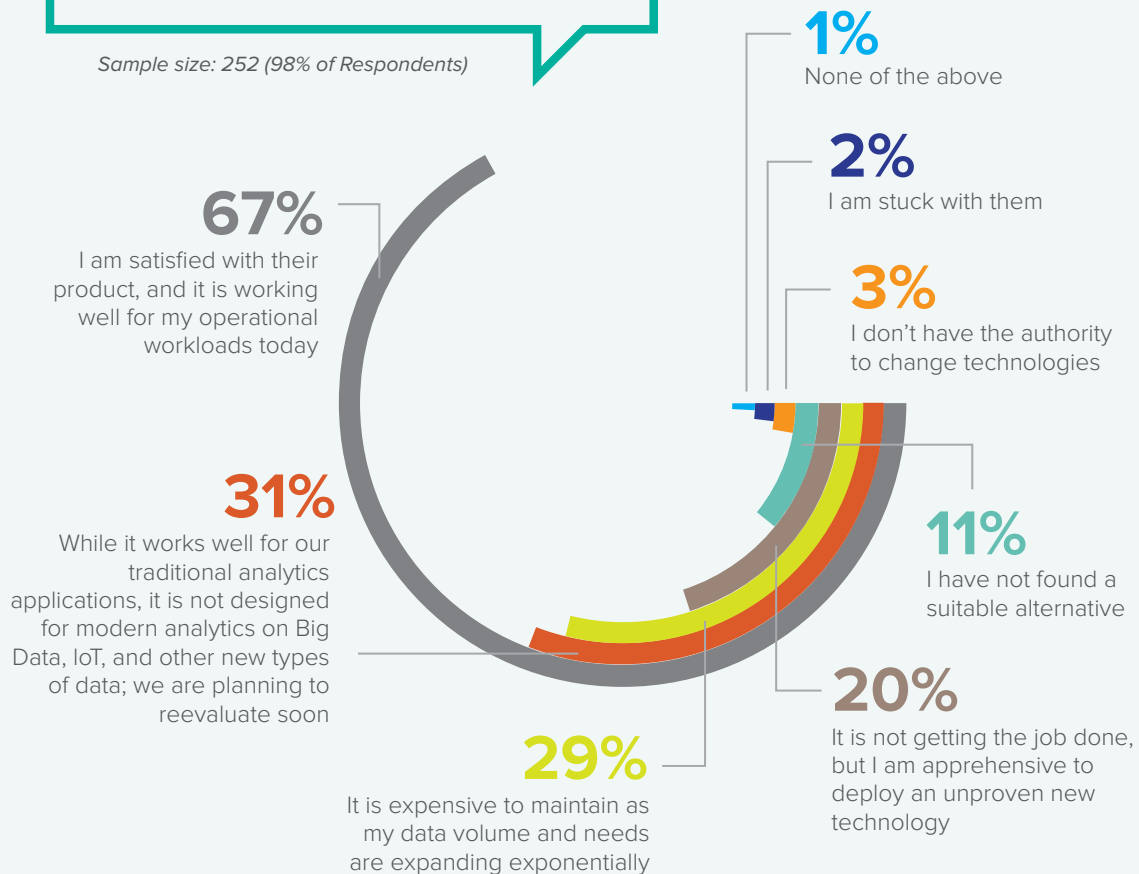
Sample size: 250 (97% of Respondents)

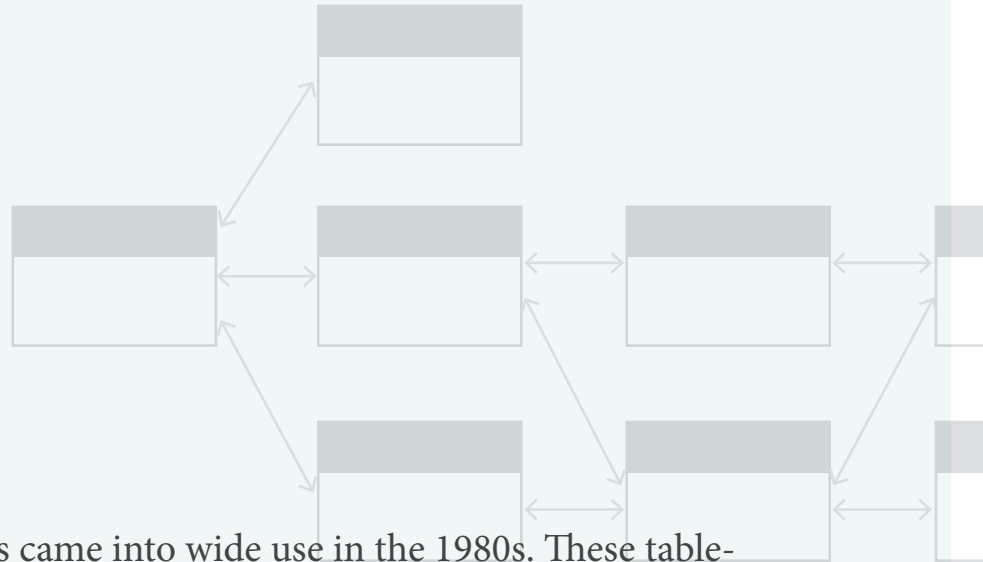


Data from the Researchscape survey also indicates that most professionals surveyed (67%) were satisfied with their current data warehousing vendor, stating that it is working well for existing operational workloads. The biggest complaint, as expected, is that it was not designed for modern analytics on Big Data, the IoT (Internet of Things), and other new types of data (31%). The second-biggest complaint is that it was expensive to maintain as data needs and volumes grow exponentially (29%).

**Which statement best represents your current relationship with your data warehousing vendor?**  
(See all that apply.)

*Sample size: 252 (98% of Respondents)*





## RDBMS

Relational databases came into wide use in the 1980s. These table-based (column and row) relational databases became popular because they provided many useful features:

- Reliability
- ACID Compliance
- Aggregation
- Rigorous Schema
- Complex Queries
- Normalization

Relational databases are used to process and store structured data, and they offer many tools, but also a significant amount of infrastructure and overhead. A relational database simply cannot meet all of the business requirements and opportunities that have arisen with big data. The biggest limitation is that you are bound to a rigid schema ahead of time.

## Dealing with Big Data



Big data is a direct result of both the explosive growth of the Internet, and the proliferation of social and mobile applications. To deal effectively with big data, a database must possess the following characteristics:

- **Scalability** – The database must be able to scale up to hundreds of terabytes of data, and scale dynamically up to thousands or millions of users.
- **Flexibility** – The database must be able to deal with schema-less data sets that include structured and semi-structured data.
- **Performance** – The database must provide fast transaction response times, at scale.
- **Agility** – The database must provide developers with the ability to rapidly create and adopt applications.

There are two ways to scale a database:



### **Scale Up**

Scale vertically by adding CPU, storage, and memory to individual servers



### **Scale Out**

Scale horizontally by adding more servers in parallel

---

Legacy relational database systems cannot scale out effectively. An RDBMS server can be scaled up by adding CPU, storage, and memory, but you eventually reach an upper limit that still cannot adequately address the performance requirements of big data.

A NoSQL database is designed to scale horizontally, leveraging the processing and storage power of hundreds or even thousands of parallel servers. And instead of requiring expensive, specialized storage and processing hardware, NoSQL takes advantage of low cost commodity servers. In addition, NoSQL is schema-agnostic which accelerates development cycles by providing the freedom to store information without requiring extensive up-front schema design.



# Why Move to NoSQL?

Using NoSQL does not necessarily involve scrapping your existing RDBMS and starting from scratch. NoSQL should be thought of as a tool that can be used to solve the new types of challenges associated with big data. There may be business processes that can continue to be addressed effectively with RDBMS. But with the new challenges presented by big data, you will likely face new problems that can be solved more efficiently – and more cost-effectively – with a NoSQL database.

A NoSQL database can be used to solve new problems that require:

- **Scalability** – A NoSQL database can scale horizontally to the scale required by big data. Applications can run in parallel on a cloud-based cluster comprising of dozens, hundreds, or even thousands of commodity servers. The NoSQL scale-out architecture enables web applications to scale dynamically up to thousands or even millions of users. As the number of concurrent users grows, you can dynamically add more cluster nodes (commodity servers) to process the additional load.
- **Flexibility** – A schema-less NoSQL database can process and store structured, unstructured and semi-structured data, and enable flexible and rapid development of applications and use cases such as right-time decisioning, recommendations, profile management, bidding, and risk profiling. With RDBMS, you cannot add data (columns) without updating the entire schema, which can take hours or even days, depending on your institutional infrastructure. With NoSQL, data can be added and updated flexibly and efficiently.
- **Speed** – The parallel processing nature of a NoSQL database, along with caching and aggregation, can provide fast (sub-millisecond) transaction response times at scale.
- **Developer productivity and agility** – A NoSQL database shortens time to market for new applications and updates because developers don't have to shoehorn data into a fixed schema. With NoSQL, applications can be rapidly prototyped, tested, and deployed into production in a cloud-based cluster.
- **Operational readiness** – In contrast to RDBMS, which typically require schema migration scripts, significant manual effort, and scheduled downtime for release upgrades, NoSQL schema migrations are relatively easy and have minimal impact on operational readiness. Maintenance windows are compressed, and the performance impact on users is far less noticeable.

# Why Use Aerospike?

Aerospike is a high-performance NoSQL database that enables high throughput with low latency, horizontal scalability, reliability, and operational efficiency.

Organizations across a wide range of industries, including ad tech, financial services, eCommerce, online gaming, and telco rely on Aerospike to power revenue-generating, mission-critical systems of engagement.

Aerospike has pioneered a hybrid RAM/SSD storage architecture that lets you achieve RAM-like performance at a lower cost by leveraging flash storage (SSDs). This innovative architecture delivers speed at scale, predictable performance, and high availability while dramatically reducing total cost of ownership. Applications that leverage Aerospike typically require as few as one-tenth the number of servers required by most other databases.

The NoSQL space is vast and populated by large and well-known companies (e.g., MongoDB, DataStax, and Couchbase) that are investing aggressively to go to market as Oracle replacement technology for non-mission-critical use cases. In contrast, Aerospike is focused on the next phase of the database market, where speed and scale offer companies a competitive advantage.

## Speed at Scale

Low TCO

Built for Flash storage

## Open source

Store hundreds of terabytes and access them at speeds measured in milliseconds or sub-milliseconds

Ability to support applications with a very large data set, which increases the accuracy of the decision engine

Cross Datacenter Replication (XDR) support

Predictable performance

## Next Steps



See how much you can save with the Aerospike TCO (Total Cost of Ownership) calculator!

[Go](#)



The [Aerospike Quick Start Program](#) is designed to help you fully leverage our fast, reliable, high-performance NoSQL key-value database for real-time workloads. This program allows you to quickly achieve the business benefits that our customers rave about.

The Aerospike Quick Start Program includes:

- Unlimited data usage of the Aerospike Enterprise Edition during a 90-day trial period
- Virtual Classroom Training:
  - 1 Developer Seat
  - 1 Administration & Operations Seat
- Cluster sizing and benchmark tool support for optimizing your implementation

[Learn more](#)

AEROSPIKE

© 2016  
Aerospike, Inc.