



Sponsored by:
Oracle Corp.

Authors:
Carl W. Olofson

December 2018

Ensuring a Fast, Reliable, and Secure Database Through Automation: Oracle Autonomous Database

IN THIS WHITE PAPER

Managing a modern enterprise-class database can be monstrously complicated. If it's a data warehouse, the index structure, partitioning plan, and various server settings must be just right to ensure good performance in handling complex queries. If it's a transactional database, availability and high performance are critical factors that can have real revenue implications. As databases grow in size and complexity, these concerns become compounded.

This white paper examines these issues, the approaches taken to address the issues, and the emergence of a database offering that promises to ensure the best possible availability and response times without requiring any human intervention at all. This offering is called the Oracle Autonomous Database. It works by using machine learning (ML) technology to continually self-tune, relieving database administrators (DBAs) of hundreds of hours of effort devoted to examining performance statistics and making adjustments. Oracle Autonomous Database is on the vanguard of a new class of database, one based on a self-tuning and self-healing database management system (DBMS), delivered as a cloud service.

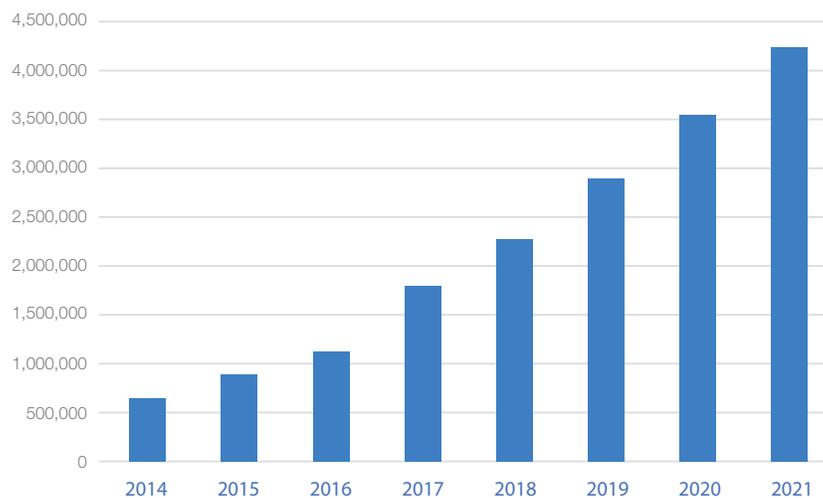
SITUATION OVERVIEW

The Database Management Challenge

Since their inception decades ago, most enterprise datacenters have grown steadily in size and complexity and today operate hundreds of production databases. Despite impressions to the contrary, relational database deployment has continued to increase. IDC expects that the

number of server systems deployed to manage relational databases will increase from 1.1 million in 2016 to 4.2 million in 2021, for a five-year compound annual growth rate of 30.2%. This finding is shown in Figure 1. These databases are applying ever-increasing pressure on both database administration and operations staffs, despite the fact that most organizations have been trimming, not expanding, those staffs.

FIGURE 1 Worldwide RDBMS Server Shipments, 2014–2021



Source: IDC, 2018

While many of these databases serve specialized or minor supporting purposes, many others serve mission-critical functions, processing data that is used to transact business, maintain fruitful customer relationships, and ensure the ongoing smooth operation of the enterprise. When these databases run slowly or worse, crash, they can slow the pace of business, impair customer relations, and cause lost business opportunities. When problems involving data warehouses arise, management can't make critical decisions and may have to "fly blind" while the problems are being addressed.

Database failures can cause a chain reaction of disfunction: In most cases, databases are tied to other databases through either streaming data connections or extract, transform, and load (ETL) processes, so a breakdown on the source end can negatively impact the target. Most datacenters have elaborate procedures in place to minimize the incidence of failure and also to minimize its impact when it does occur. These procedures are largely manual.

Adding to these issues is the fact that although enterprises keep adding applications and their databases, and integrating that data, as well as data from new sources, such as data lakes

and external streaming sources, into data warehouses of ever greater complexity, budget constraints prevent the increase in staff necessary to deal with the additional work. As a consequence, the overall quality of operations suffers. Not only does this negatively impact existing functionality, but it inhibits the implementation of new functionality; the database administrator who would set up a new database workload must put off that job because existing databases require so much attention, and there is only so much time in the day.

Threats to the Business Posed by Poor Database Management

When databases are not maintained and tuned fully, problems that represent material risks to the business can arise. These problems include interruptions or delays in data handling that interfere with business transactions and slow down proper decision making.

The problems can also include data security issues, since DBMS software vendors are constantly issuing patches to address vulnerabilities. When those patches are not applied in a timely manner, the risk of a data breach looms. Because applying patches manually in the datacenter can be operationally disruptive, and requires extra time and effort, such patches are often batched up, with the result that the database is left at risk for some period of time.

In a large database, the composition of data is in a constant state of flux, requiring constant analysis and tuning to ensure good performance. Failure to do this can result in sluggish performance and even outages if the internal database resources are insufficient to handle the volume of data being processed. Slow or unavailable databases can result in slower transaction rates, less business being done, and lost opportunities.

The problem may be even more acute for analytic databases, such as data warehouses. These are increasingly used to provide a complete view not only of the business but of the business environment, including customers, prospects, competitors, and various secular risks and threats. For this analytic data to be complete, accurate, and current, the database must be well tuned and constantly adjusted to the shifting requirements of the data being analyzed. This normally requires an enormous ongoing manual effort to ensure not only that the database performs well but that it can be made available to all business users who need access to the data in order to make critical business decisions.

Manually Managed Databases: Costs, Risks, and Limitations

The single largest inhibitor to ideal database configuration is the number of staff that handles the tuning, problem determination and correction, and routine operations (software patches, system maintenance, etc.). Although the databases grow constantly in size and complexity, the size of the staff tends to remain constant. Yet the DBA staff is under constant pressure

to maintain excellent performance while constantly adapting the database to the changing needs of applications and users. Other requirements, such as changes to data organization for application or analytic purposes, often must wait as performance and reliability issues always come first.

In general, the costs and risks associated with enterprise databases may be summarized in the following ways:

- **Infrastructure cost:** This is the cost of acquiring and maintaining the servers, storage, and networking infrastructure required to run the databases. Inefficient database operation can cause excessive expenditures in this area, as implementers attempt to “throw hardware at the problem.”
- **IT staff productivity:** IT staff, including the DBA staff, must work to make the system with as much performance as possible. The operations staff’s work is fairly routine, but the DBA’s work is complex, requiring much analysis and expert handling of the databases involved. Since DBAs are also needed to implement new databases for new applications or change databases to meet the needs of changing applications, overworking the DBA team can result in applications that are simply not responsive to the changing needs of the enterprise.
- **User productivity:** If users are exposed to excessive wait times or, worse, to application failure, they can’t get their own work done. This results in extra cost to the business if the user is internal and lost business if the user is external.
- **Business productivity:** When applications don’t meet the needs of the enterprise, when business analysts can’t get the data they need to make key decisions, when employees can’t do their jobs, and when customers are frustrated in their efforts to interact with the system, the result is lost business opportunity and lower revenue.
- **Other areas:** It has often been said that “data is the lifeblood of business.” If the flow of data is slowed or interrupted, there can be other negative consequences that may be hard to track or quantify but nonetheless result in a less efficient and less profitable business.

Business Requirements

Businesses need to ensure that as the database evolves, it continues to adjust optimally to the growing size and complexity of the data. Ideally, as it grows the database must continue to deliver the following capabilities:

- Exhibit excellent performance, reliability, and availability.

- Enable business users to innovate at the speed of business.
- Enable users to have ready access to the latest data in the most useful and relevant forms.

One way to do this is to move database operations to a public cloud service. This transfers the effort associated with ordinary operations from the datacenter IT staff to the cloud service, which can offer guarantees regarding availability and overall system performance. Such guarantees, however, do not necessarily extend to the database. Since the user organization defines the database, the user organization assumes responsibility for its smooth operation, so the larger locus of risk, which is in the work of the operational DBA, persists.

The ideal solution is for the database to be managed as a service, which entirely relieves the responsibility of the DBA staff for tuning and troubleshooting the database. But how is this possible, if the responsibility for overall database definition rests with the user? What is required is technology that can tune and recover a database without human intervention. This is a tall order, since there are many factors involved in these activities, and even experienced and expert DBAs struggle to carry them out. The DBMS can't simply execute rules or scripts. It must actually adapt as the database changes and must learn from successive actions and their consequences.

The system that can do this not only ensures that critical enterprise data is always available, and so operational applications and analytics can work reliably on the latest data, but also enables DBAs to help the database applications development teams and business analysts adapt the system to the ever-changing requirements of the enterprise.

Introducing the Autonomic Class of Database

To make the delivery of a fully self-managing database-as-a-cloud service a reality, one must come to terms with the complexity of the problem. A database cloud service that is manually administered is likely to disappoint when it comes to such things as ensuring scaling, performance, and availability to users at an optimal level; correcting faults as they occur; and anticipating problems before they occur. This is because a large, complex database, such as a data warehouse, cannot be so administered without incurring such a high cost to the service provider that the fee would be unacceptably high.

The answer is to make the database itself self-tuning and self-healing, adapting to changing conditions as they arise. This approach involves autonomic computing. It's not a new idea. Autonomic computing has been applied to problems of systems management and general operations management for some time. It has generally been found inadequate for large, complex systems because, usually, it has been deployed algorithmically, that is, program code or sets or rules determine responses to various situations. Such a technology can, of course,

only respond effectively to problems previously seen, defined, and accounted for in the coding. Autonomic database technology such as this has been offered by Oracle and other RDBMS vendors for some time.

The application of autonomic computing to complex systems like data warehouses requires more than algorithmic, rules-driven self-operation. It requires that those systems be able to learn from existing conditions and adapt to them. This involves artificial intelligence driven by machine learning. Oracle has established an early lead in this area by introducing the Oracle Autonomous Database.

Oracle Autonomous Database

Oracle Autonomous Database uses ML technology at several levels:

- Database tuning, for optimal performance
- Database error correction and recovery, for optimal uptime
- Database security, for detecting anomalous or suspicious usage in real time

Oracle Autonomous Database is available for the optimization of several usage models: data warehouse, enterprise OLTP and mixed workload (including analytic-transaction processing), and development databases. This system also features elastic resource utilization, growing and shrinking processor and storage usage (within user-defined limits) as database size and user demand vary, so the system is always the right size and the user only pays for the resources necessary to run the database in an optimal manner.

The user controls only the schema and the data itself. As these change, Oracle Autonomous Database adjusts its setting for optimal performance and operation. Oracle Autonomous Database also takes advantage of Oracle's high-availability (HA) capabilities for continuous operation and optimizes routine operational functions such as backup and patching. As a consequence, Oracle can offer a 99.995% availability guarantee, with, as Oracle puts it, "no ridiculous exclusions."

This capability also dramatically reduces the amount of staff time required by the database. The lead DBA for a major personal care products maker reckons that the Autonomous Database would save his organization 10-30 percent staff time savings, which means that staff effort, on average, would be reduced from full time to 10-20 hours per week. This is less about cost savings, and more about having a more effective team. The resulting staff redeployment promises to make the group more flexible and efficient.

The automatic patching feature of this service represents a major benefit as well, ensuring that security fixes are applied as soon as they are available, without the cost, disruption, and risk associated with applying patches in databases running in the datacenter. For QLX, a supplier of data services to sports enterprises and casinos, security is paramount. The security capabilities of Oracle Autonomous Data Warehouse (Oracle ADW), and in particular the automatic patching feature, makes it a huge win. QLX clients often manage venues that are tied into Homeland Security operations. They have a high level of requirement with respect to access, authentication, and data encapsulation. A hack would be disastrous. This feature makes such an eventuality much less likely.

QLX also found Oracle ADW to perform 10 times better than its on-prem data warehouses. They ran tests involving machine learning, SAS, Python modules and the Oracle Analytics Cloud, and found the same performance benefit in all cases

Oracle Autonomous Database Warehouse

Oracle Autonomous Database Warehouse is the first of a series of cloud services based on the Oracle Autonomous Database technology. It is a database cloud service configured to operate as a self-managing, self-tuning, and adaptive data warehouse. This cloud service is offered either as a database as a service (DBaaS) in the Oracle Cloud or in the datacenter through Oracle's localized cloud service appliance, Oracle Cloud at Customer. It operates in an environment that is entirely under Oracle's control. In this way, the service can provide optimal performance for the workload at hand. The user does not specify resources such as processors or storage; rather, the Oracle Autonomous Database selects a configuration that is ideal for the database that it is managing. Oracle Autonomous Database also makes all decisions regarding such things as indexing, partitioning, and the use of Oracle's hybrid columnar compression.

Oracle Autonomous Database is designed to support both structured and unstructured data. Oracle Autonomous Database Warehouse (ORACLE ADW) designed for data warehousing use cases and Autonomous Transaction Processing (ATP) supports transactional databases and mixed workloads.

Benefits

Oracle promises the following benefits from using Oracle Autonomous Database:

- Flexible, scalable operations ensure that the users, including business analysts, can get consistently excellent performance reliably on even the largest, most complex, and changeable data warehouses.

- Continuous operation, even when maintenance is being performed or the operating parameters change. There are no interruptions in service as patches are applied, or when the database is being scaled up or down in terms either of storage or processors.
- Optimal data security is ensured by virtue of the fact that the database software is always up to date with the latest security patches that are designed to prevent unauthorized data access.
- Elimination of staff time spent on DB operational administration allows DBAs to focus on the more high-value functions of working with application developers and business analysts to ensure that each Oracle Database has the optimal data required for the enterprise to operate in and adapt to ever-changing business conditions.
- With Oracle Autonomous Database, as more databases are added, there is no additional operational staff time cost because Oracle Autonomous Database performs those functions itself.
- Barring a massive multisite disaster, a database under Oracle Autonomous Database's control should never fail, virtually eliminating downtime.
- The ML capabilities of Oracle Autonomous Database are constantly learning and adapting as the database runs to ensure the best possible performance of the engine and of queries.
- Sophisticated data protection is based on the security capabilities of the Oracle Autonomous Database and its ML features that discover and prevent inappropriate data access.

User Experience

Minnesota State University

Minnesota State is the 5th largest institution of higher education in the country. The university is a longtime Oracle user on-prem and was an early beta tester for the Oracle Autonomous Data Warehouse. According to Minnesota State's systems architect, Miles Oustad, the school was an active participant, providing input regarding functionality as the process went along. Minnesota is "cloud-first" in its forward-looking IT strategy, so their interest in this service was natural.

The testing looked at levels of deployment, data upload speed, and performance testing of typical data warehouse queries. They built a test dataset from the production system, filtering out any personal identifying information (PII). The work was done by the data warehousing team, which Mr. Oustad described as being more like data scientists than DBAs because of their intimate knowledge of the data and the analytics models.

His observations included the following:

- They were able to provision the system in minutes; a key capability.
- They did not need to concern themselves with physical details of the setup, and the usually arduous DBA task list for such setup had disappeared.
- Two full time people are freed from doing any physical DBA work by this service.
- Table scans ran 100 times faster than the same scans on their on-prem system.
- Complex combinations of data warehouse queries, representing a typical workload, improved in throughput by 95 times.
- The load time was just fine.

Oustad identified the following as key benefits:

- The big labor cost savings in system management and tuning.
- Relief from having to buy future capacity today, replaced by a pay-as-you-grow model.
- No more delays, complexity, and risk associated with manual system patching.

Oustad concluded by saying that, both functionally and operationally, Oracle Autonomous Data Warehouse is “an awesome way to go.”

FUTURE OUTLOOK

The future of most enterprise data management is in the cloud; this is pretty much a consensus view. The question is, how will that management be done? While cloud services teams can perform routine tasks such as taking backups, applying software patches, and tuning the systems on which the database runs, actually troubleshooting database performance problems or failures may require a level of intervention that is a little too revealing of the enterprise database to the cloud services staff. Also, it is a non-scalable area of work that is likely to contribute to a high service cost for DBaaS.

So, it seems clear that a successful DBaaS must include autonomic features in the DBMS itself, including self-tuning and self-healing operations driven by ML. Oracle Autonomous Database makes Oracle a pioneer in this regard and gives Oracle Cloud a distinct early advantage.

CHALLENGES/OPPORTUNITIES

There can be no doubt that other database software vendors will develop autonomic features similar to those in the Oracle Autonomous Database. At the same time, Oracle has the opportunity to leverage this capability as long as the firm continues to innovate along these lines, offering capabilities that not only match the abilities of the most talented and experienced DBAs but go beyond to provide tuning and recovery capabilities that no human can hope to match.

CONCLUSION

Enterprise databases are getting more numerous and more complex all the time. Their interdependencies through data integration also sets them up to fall like dominoes in cases of sluggishness or outright failure. To prevent this, enterprises employ armies of operational DBAs to continually monitor, tune, and recover databases. Yet budget constraints often prevent those DBA teams from growing to match the work required, so the higher-value work of making databases more responsive to the needs of the enterprise often fall by the wayside.

Public cloud services offer a partial response to this problem by alleviating the routine work of system, application, and database maintenance, but the work of keeping the databases running with peak performance is still a labor-intensive function. To address this problem, enterprises should consider the following:

- Make a comprehensive plan for moving all systems not requiring locally resident data to a public cloud service.
- Consider a locally delivered cloud service appliance, such as Oracle's Oracle Cloud@Customer, for the management of data that must remain local.
- Tally the costs and risks that are associated with the current manual method of managing enterprise databases in your datacenter. If those costs and risks are unacceptable, consider turning to Oracle Autonomous Database as a means of protecting the enterprise from unnecessary constraints on enterprise performance, productivity, and agility.

IDC Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com

Copyright Notice

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

Copyright 2018 IDC.

Reproduction without written permission is completely forbidden.

About IDC

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