



The Evolving Role of DBAs in the Cloud

Examining the Challenging of Managing Databases in IaaS, DBaaS and Hybrid Environments

AUTHOR

Warner Chaves

Principal Consultant, Microsoft Certified Master, Microsoft MVP

CONTRIBUTOR

Chris Presley

Director, Solutions Architecture



Soaring cloud services adoption is causing a seismic shift in the role of DBAs. Whether an organization chooses infrastructure-as-a-service (IaaS), database-as-a-service (DBaaS) or a hybrid architecture, the role of DBAs changes significantly. This white paper examines the evolving role of DBAs and discusses how database managed services, or DBA services, fit into this new landscape of managing databases in the public cloud.

The View from the Cloud: A Changing Landscape for DBAs

The accelerating adoption of cloud services continues unabated, with research firm Gartner **observing** that cloud computing is now “firmly established as the ‘new normal’ for enterprise IT” and “one of the fastest-growing segments of IT spend.” The numbers bear this out. With cloud revenue growing at an average clip of more than 40 percent each year since 2011, Gartner **also says** infrastructure-as-a-service (IaaS) spending will almost double between 2019 and 2021, from \$39.5 billion to \$63 billion. At the same time it says data center spending will fall to \$190 billion through 2022 (from \$195 billion in 2019).

And while competing research firm Forrester **maintains** that IaaS, platform-as-a-service (PaaS) and software-as-a-service (SaaS) revenue will likely begin to slow in the next couple of years, they will still account for a combined \$411 billion, with a strong compound annual growth rate (CAGR) of 21 percent by 2022.

Indeed, according to a **Pythian-Unisphere research report from 2019**, 73 percent of organizations surveyed have plans to deploy cloud-native data lakes within two years. The same report indicates that 48 percent of organizations already have cloud or hybrid data warehouses, with another 48 percent planning on migrating to the cloud within the next few years.

The ideas presented in this white paper are most relevant to organizations with classic relational databases. However, they also apply to NoSQL, big data and data warehouse offerings in the cloud.

Although we provide examples from Amazon, Microsoft and Google, this paper is not a cloud vendor comparison. The vision and advice apply to all of the major public cloud service providers.

After reading this paper, you’ll have a high-level understanding of how the role of DBAs changes with IaaS, DBaaS and hybrid architectures. You’ll also understand the role of database managed services. With this information, you can better plan your cloud migration strategy and DBA resource/training needs.



This paper is targeted to IT professionals with a good understanding of databases along with business people who want an overview of changing DBA roles and want to understand the key considerations of managing databases in the cloud. We assume familiarity with the common tasks DBAs currently perform on-premise, either on hardware or virtual machines (VMs).

IaaS

IaaS Defined

Before discussing how database administration is done when using IaaS, let's define the term. IaaS is a cloud service model in which the virtual data center, its physical location and the physical hardware are managed by the service provider.

The DBA maintains full responsibility for the software that is inside the VM. This includes the operating system (OS), any utilities on the OS and the database software. The DBA chooses and manages all of these. The DBA can scale up or scale down a particular VM on demand and can also create or destroy VMs on demand.

The core tasks of DBAs do not change with IaaS. What changes is how DBAs perform these tasks.

Capacity Planning

Capacity planning has been a core DBA task since the release of the first database engines. Performing this planning in IaaS is very similar to doing it on-premise for a virtualized environment. The key difference is the added flexibility due to the fact that cloud providers will not run out of physical VM hosts. As a result, DBAs should provision for only what the business needs today instead of old-style over-provisioning for a physical machine that is expected to last three to four years.

DBAs need to make use of the bursting capability of the cloud to quickly scale up or scale down to meet demand. Proper capacity planning also involves responding to changes in offerings by cloud service providers to get the best deals. For example, a provider may offer a new type of storage or a new type of VM that would be a better fit for one of your workloads.

In this new world of cloud IaaS, DBAs need to be highly aware of the ebb and flow of demand. They need to collaborate with the business to prepare for surge events. Because on-demand scalability is a key benefit of cloud platforms, the modern DBA needs to understand not only how to make use of this feature, but how to anticipate and respond to major events or cyclical demands on their databases, then spin



their capacity up or down as needed. Being responsive, and increasing capacity only when needed, is critical to controlling costs. For example, events demanding short-term increases in capacity include Black Friday in the retail sector, concerts or major events for ticket suppliers and the start of the school year for educational websites.

Monitoring

Monitoring is another core task that DBAs perform daily. All of the major cloud IaaS providers offer monitoring. Amazon Web Services (AWS) has a built-in service called Amazon CloudWatch that monitors any type of AWS asset. Google offers Stackdriver Monitoring and logging that integrates with all services provided by Google Cloud Platform, including Cloud SQL. Microsoft Azure provides the Azure Monitor service, which enables DBAs to put different thresholds on different metrics tied to a service, provides alerts dependent on those values and can be used for analysis over time.

These services are very similar in concept but different in implementation. As a result, DBAs require in-depth knowledge of how to configure and test monitoring in whichever IaaS offering they are working with.

The monitoring tools provided by the IaaS providers are a start. However, DBAs will often want to augment these with more specific monitoring tools. Use cases include monitoring for key metrics of the business application or monitoring the application end-to-end. It's important to understand how to build custom monitoring in whatever cloud platform you are managing.

High Availability and Disaster Recovery

Mission-critical databases inherently have very high up-time requirements that entail the need for high availability (HA) and disaster recovery (DR) architectures and strategies. This is still true for IaaS in the cloud.

To provide proper HA and DR for IaaS, DBAs need to know which clustering technologies are the most cloud friendly, for example, if it's better or easier to have a share-it-all cluster or to instead have a local storage cluster. DBAs also need to know how to architect the HA and DR and how to reconfigure them on IaaS. The concepts of HA and DR remain the same for all the major public cloud IaaS offerings but the physical implementation differs quite a bit from service to service.

AWS offers the concept of regions, and inside the regions there are availability zones. This allows the infrastructure in one availability zone to fail over to infrastructure in a different availability zone.



Like AWS, Microsoft Azure has the concept of paired regions. Azure automatically replicates some of its assets to a prespecified paired region. Azure also offers Availability Zones inside one region as well as the concept of availability sets; if you put two VMs in one availability set, they'll provide high availability between them.

In addition, Azure has fault domains whereby the VMs are placed on different physical racks, and also update domains. This architecture ensures that when any type of software update is performed on the underlying infrastructure, two VMs are not updated at the same time.

Similarly, Google delivers services across many regions covering North and South America, Western and Eastern Europe, Asia, and Australia. Each region is represented by multiple zones, allowing you to distribute your apps and storage across to protect against service disruptions. It's possible to make use of zones in different regions for additional redundancy.

Configuration and Performance

In cloud IaaS, configuration optimization and performance tuning also change. On-premises, a business usually has just one hardware provider. The DBA changes the hardware settings as required to optimize configuration and improve performance. By contrast, each IaaS provider has a different family of VM types for different database workloads.

In AWS it's common for people to deploy database workloads on general-purpose VMs, or on memory-optimized VMs if they're dealing with a high-throughput, low-latency application.

Inside Google Cloud Platform you can choose Cloud SQL using either standard or high-memory machines based on application demands.

In Microsoft Azure, small to medium database workloads are commonly deployed on general purpose VMs and larger higher performance workloads on memory-optimized VMs that offer high RAM/core ratios.

Another IaaS configuration consideration is storage—which is different for each service provider. AWS offers general SSD storage and provisioned IOPS SSD storage, which is designed for I/O-intensive workloads. These two are the most common storage options deployed for database workloads. There are also different types of magnetic storage that you can add to a VM if this option fits well with your architecture and solution.



Microsoft Azure provides magnetic, standard SSD, premium SSD and ultra disk storage on different tiers that support different sizes and performance characteristics from low-volume cost effectiveness to super-high-throughput, low-latency response. It's important for DBAs to understand the capacity of any of those tiers, and how to mix them to get the right combination of performance and cost-effectiveness.

Without the proper configuration optimization and performance tuning, a DBA can run into two unpleasant situations: not getting the best performance from a particular VM, or paying more for a VM that delivers less. For example, improper storage configuration and tuning paired with the wrong VM instance can result in paying more for storage that isn't performing to 100 percent of its potential.

Summary

If an organization is moving to cloud IaaS, the DBA needs to understand how all the core database administration tasks are done differently with IaaS—and also how they are done differently on each IaaS offering. With this knowledge, the DBA can advise on the best choice for the organization. Without this knowledge, businesses risk becoming locked into a particular cloud provider because of the lack of in-house skills.

Almost 50% of organizations in [one study](#) pointed to a lack of skills as their top challenge in implementing cloud-based infrastructures.

A trusted advisor, whether an in-house resource or an IT service provider, will also be on the lookout for efficiency opportunities. For example, moving a database that is now on-premise to a VM to a cloud DBaaS may reduce your costs and management—and that's the main reason to move to the cloud.

DBaaS

DBaaS Defined

Before discussing how database administration is done when using DBaaS, let's define the term. DBaaS is a service which provides the underlying infrastructure and RDBMS capabilities in a fully managed and elastic offering priced by utilization. DBaaS unburdens the DBA from having to manage infrastructure, operating systems or RDBMS patching and maintenance. In addition, high availability and data protection is inherently built in. Scalability is elastic and there's no need to over provision in anticipation of future growth.

How DBA Roles and Tasks Change

With DBaaS, the core tasks of a DBA do change—because DBaaS aims to automate most, and eventually all, operational tasks (see Table 1). As a result, DBAs with a skill set that is 100 percent based on operational tasks need to broaden their skill set or risk becoming obsolete as more businesses migrate to DBaaS offerings to decrease their management footprint.

Table 1: DBA Roles and Tasks with IaaS and DBaaS

DBA role with IaaS	DBA role with DBaaS
Manage the OS	No access to the OS
Install patches	No patching
Upgrade the system	One-button upgrade
Perform backup and recovery	Backup and recovery performed by service provider
Determine and implement HA and DR architecture and strategies	HA and DR are inherently built in

The DBA role can now be refined to operational tasks rather than infrastructure maintenance and support. The DBA's skills will still be required for schema evolution, data management, data ingestion, user and RBAC role management and administration, and data recovery due to operational loss, i.e. inadvertent dropping of a table. The DBA can now focus primarily on query performance, meeting business objectives, and less on maintaining the platform.

Performance Tuning

Performance tuning also changes with DBaaS. A DBA needs to know the mechanisms that the service provider makes available to increase or decrease the compute power, and have the knowledge to choose the most effective and cost-efficient solution.

To increase or decrease compute power in Amazon Relational Database Service (RDS), a DBA can change the instance type of an RDS database server. In Google Cloud Platform, the DBA can change the instance type for Google Cloud SQL deployment. And for Microsoft Azure, the DBA can increase or decrease the database transaction units or use a core-based sizing model.

On low-usage or burst-based database workloads, a serverless database might make more sense, where you pay specifically based on CPU consumed and you don't reserve resources at any time.



As new features are released every month in the cloud, it's also important for DBAs to be knowledgeable about them and provide valuable recommendations on which ones can be used to increase the performance of a business's databases. Examples of new features include columnar storage and read-only replicas.

Performance tuning is a constant, iterative process. It may be easiest to just ask the DBaaS provider for more compute power—but this could result in a bill increase. This cost is often avoidable if the DBA is an experienced database tuner.

The major public cloud vendors offer advanced monitoring and alerting functions which normally require substantial and expensive add-ons to on-premise RDBMSs, i.e. Oracle Enterprise Manager, and thus require their own RDBMS which in turn requires maintenance and patching, adding to the DBA's workload. Public cloud-based monitoring and alerting provides a much more granular level of detail, and provides the added benefit of usage-based cost reporting.

Custom Monitoring

As applications grow in complexity and user expectations keep getting higher, monitoring only generic performance levels is simply not enough. DBAs need to be comfortable using various monitoring systems—not just those offered by the main DBaaS providers, but also those from third-party tool vendors.

DBAs also need to be able to develop scripts for very customized alerts to signal issues that are specific to their application, for example, developing an alert for an order queuing table that is growing out of control.

Automation Through Scripting

Automation is another core task that must be adopted by DBAs in this cloud age. The DBaaS providers offer functionality to automate deployment of objects. AWS has CloudFormation templates. Google has Google Cloud Deployment Manager. Microsoft Azure has Azure Resource Manager templates. These are all the same concept, but with different implementations. Automation templates are just a starting point. There are also popular third-party options like Ansible and Terraform.

DBAs need to help their organizations standardize these templates and ensure that best practices are defined for the database tier, with processes that are repeatable and reliable. In addition, DBAs need to work with application owners to respond quickly to application and development requirements.



Security

Security is a major concern in the cloud, and it will continue to be a DBA responsibility. DBAs need to be familiar with which parts of their databases hold sensitive data, including any type of personal identifiable information. DBAs also need to know the change control and procedure measures that must be followed to provide access to sensitive database resources. Finally, they need to be aware of new security features released by DBaaS providers, to determine which features can be used to improve the organization's security, including encryption and protection of its databases.

Summary

DBaaS significantly changes the role of DBAs. Core operational tasks are automated or managed by the service provider. As a result, DBAs whose skill set is 100 percent based on operational tasks need to broaden their skill set or risk becoming obsolete.

The new role of DBAs involves code review, performance tuning, custom monitoring, automation and security. In all these areas, DBAs can add value through their deep knowledge of the organization's database. But they will also need to understand all the major cloud DBaaS offerings in order to make the best choices for their organization.

Even more so than with cloud IaaS, a trusted advisor—either an in-house resource or IT service provider—will also be on the lookout for efficiencies.

Hybrid Architectures

Because most businesses will still keep some components on-premises, another model of cloud adoption is the building of hybrid architectures for both IaaS and DBaaS. According to IDC, more than 70 percent of heavy cloud users are considering a hybrid cloud strategy.

To determine when a hybrid approach is a good fit, DBAs need to have experience in implementing and monitoring both IaaS and DBaaS and dealing with all the issues previously discussed for both. Furthermore, DBAs require in-depth knowledge of the differences in implementation and monitoring with each of the main IaaS providers and also each of the main DBaaS providers. DBAs also need to know the pros and cons of running on-premises versus running on each of the main public cloud providers for both IaaS and DBaaS.

Only with this depth of knowledge can DBAs decide when the hybrid approach will provide the best return on investment and also which cloud provider to choose for the hybrid architecture. This is a daunting amount of specialized knowledge to expect to find in a single DBA.



Data-centric Development

Looking further into the future, DBAs need to have skills for data-centric development; this includes automation through scripting, as already discussed. In addition, data warehousing and big data are now more common. The cloud has made it easier for small and medium-sized businesses to adopt these types of offerings and gain competitive advantages from them.

Architecture for, and delivery of, reporting and analytics go hand in hand with the increase in data warehousing and big data. As a result, DBAs will need to know how to architect for reporting and analytics and how to deliver the reporting and analytics data.

DBAs will also need to know about data integration between systems because today's businesses take data from the web, from on-premises and from partners and providers, then try to integrate it to gain insights. Streaming and real-time data flows are also becoming more important because businesses require more insight—and faster insight—into what's happening with their operations.

Finally, the field of data science has become increasingly important. The role of DBAs is to operationalize the work of data scientists so that it can be deployed in a highly available and high-performing manner in the cloud.

The Role of Expert Database Services

Before we discuss the role of database managed services and their relationship to the new, evolved role of DBAs, let's take a step back—because the role of expert technical services is directly proportional to the importance of databases themselves.

Why Databases are so Important

Relational databases were developed at a time when storage and memory were extremely expensive. The databases' primary purpose was the ability to retrieve data quickly and efficiently. The major technological innovations at the time were the “least recently used” algorithm, which cached frequently used data to memory, and block-based storage, which increased the efficiency of data reads and writes to/from disk. Transactional integrity was a key requirement for financial transactions and provided “read consistency” to assure a consistent view of committed data, which was provided by advanced row-locking schemes. Finally, the introduction of SQL provided programmers a relatively easy way to access the data.

Today's modern cloud-based storage and retrieval systems, e.g. NoSQL, aren't burdened by past limitations and have virtually unlimited



resources. While transactional integrity is still best served by an RDBMS, data warehouses and analytics/reporting can be better achieved with nonrelational data storage such as NoSQL databases, Parquet file formats and Big Data processing.

The DBA role will now be required to evolve to understand not just how data is stored and accessed, but how to automate provisioning through DevOps practices such as infrastructure as code (IaC) and data pipeline development and management. There will be a shift from infrastructure management to data management.

In this new landscape, DBAs need to understand the differences in how IaaS and DBaaS are architected, executed for production, and properly managed and monitored. DBAs need to understand all of these things for each of the main cloud providers of both IaaS and DBaaS. DBAs also need to understand how to build hybrid architectures—and when that makes sense. Not understanding how these services work, and how to leverage them properly, can lead to costly incidents. Looking forward, DBAs also need to develop the skills for data-centric development.

The Value of Outsourcing Expert Database Services

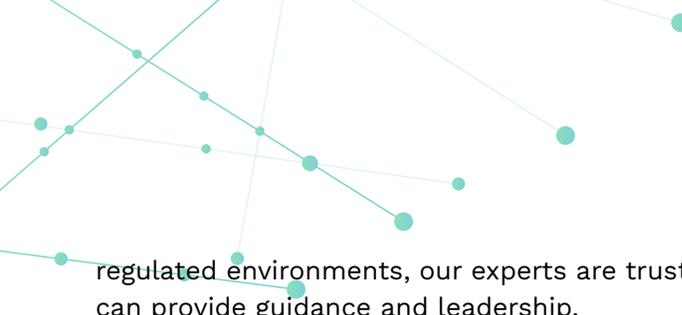
The challenges stemming from an on-premise to cloud native migration are, at first glance, daunting. As explained earlier, the role of the DBA and other IT staff will dramatically change in nature and scope. Few organizations are prepared to make the transition, and this fear is what delays a successful migration.

What's needed is a team with expertise in both legacy data center environments and cloud native technologies, and with the design patterns, tools and experience to guide organizations through this transition. A proper analysis of the business's IT current state and a thorough review of the desired future state can be provided by professional services organizations specializing in enterprise migrations to the cloud.

Why Pythian for Expert Database Services

Pythian provides you with technology-enabled experts that work in dedicated, global teams. Our secure delivery of services is the best in the industry and our contracts are flexible, so you aren't locked in long term.

Our elite experts help you use technology for competitive advantage. They bring value to your business through highly specialized knowledge: they're Oracle Aces, Microsoft MVPs, cloud, big data and open-source experts. With experience in all industry sectors and in companies ranging from start-ups to multi-billion-dollar corporations in highly



regulated environments, our experts are trusted strategic advisors who can provide guidance and leadership.

Our dedicated teams are available globally, 24 hours a day, seven days a week. Each team handles a small number of clients, so we become intimately familiar with your systems and your in-house resources. We recognize the value of your internal staff, and we can work as an extension of your team, sharing and learning from each other.

Top talent, mature methodologies and industry-leading tools provide peace of mind. Our advanced, secure service delivery processes provide the industry's highest standard of care for your business.

Flexible, month-to-month contracts guarantee that you're not locked in long term. You can easily increase, reduce or cancel your contract with only 30 days' notice, and experience the Pythian difference risk-free.

Are you ready for change? Pythian Database Services help clients not only survive, but thrive, in today's rapidly changing database environment. For more information about how Pythian can provide DBA services in the cloud, go to www.pythian.com/database-services.

Conclusion

Adoption of public cloud services continues to increase, with ever more businesses deciding to adopt IaaS, DBaaS or a hybrid architecture. As a result, the role of DBAs has changed significantly. And the advent of data-centric development promises yet more changes to come.

To choose the most efficient and cost-effective option for your organization, the DBA needs to understand the pros and cons of IaaS, DBaaS and hybrid architectures—and also needs to understand the differences among the major public IaaS and DBaaS cloud services.

After you've moved to the cloud, your mission-critical systems still require a high level of expertise to avoid costly mistakes. In addition, the cloud has accelerated the rate of change and innovation. The only way to not just survive, but thrive, is through a team-based approach.

You'll need to decide whether to augment your existing in-house DBA staff or adopt database managed services—because your mission-critical systems can't afford anything less than the highest quality of care.

For information about how Pythian can provide this care through DBA services in the cloud, go to www.pythian.com/database-services.



AUTHOR

Warner Chaves

Principal Consultant, Microsoft Certified Master, Microsoft MVP

Warner is a SQL Server MCM, Data Platform MVP and Principal Consultant at Pythian. He started his career with a brief stint in .NET programming that led to working for enterprise customers in Hewlett-Packard ITO organization. From there he transitioned to his current position at Pythian, building and managing data solutions for some of the biggest names across many industry verticals. Warner is a frequent speaker at SQLSaturday events around the world and at the PASS Summit, the largest conference dedicated to the Microsoft Data Platform.



AUTHOR

Chris Presley

Director, Solutions Architecture

Chris Presley is the Director, Solutions Architecture at Pythian. He loves order—making him a premier Microsoft SQL Server expert. Not only has he programmed and administered SQL Server, but he has also shared his expertise and passion with budding DBAs as SQL Server instructor at Conestoga College in Kitchener, Ontario. Drawing on his strong disaster-recovery skills, he monitors production environments to swiftly detect and resolve problems before they arise.



[linkedin.com/company/pythian](https://www.linkedin.com/company/pythian)



twitter.com/Pythian



Contact us at +1-866-798-4426 or info@pythian.com

ABOUT PYTHIAN

Founded in 1997, Pythian is a global IT services company that helps organizations transform how they compete and win by helping them turn data into valuable insights, predictions and products. From cloud automation to machine learning, Pythian designs, implements and supports customized solutions to the toughest data challenges.

© Pythian Services Inc. 2021

OFFICES

Ottawa, Canada New York City, USA Minneapolis, USA London, England Hyderabad, India