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

Adoption of cloud services continues to increase worldwide. In fact, Gartner says that the Infrastructure as a Service (IaaS) market has been growing by more than 40 percent in revenue per year since 2011, and it’s projected to continue to increase by more than 25 percent per year through 2019\(^1\). And according to Forrester predictions, the cloud market will grow even faster in 2017\(^2\).

Soaring cloud services adoption is causing a seismic shift in the role of DBAs (Database Administrators). Whether your organization adopts Infrastructure-as-a-Service (IaaS), Database-as-a-Service (DBaaS), or a hybrid architecture, the role of your DBAs will need to change significantly. This white paper examines the evolving role of DBAs and discusses how expert database services fit into this new landscape of managing databases in the cloud.

The ideas presented here 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 provisioning 
for a physical machine that is expected to last three to four years.

DBAs also 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 in order 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. Examples 
of 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 
Cloud Stackdriver Monitoring and logging that integrates with all services provided 
by Google Cloud Platform, including Cloud SQL. Microsoft Azure provides Azure 
Metrics Alerts, which enables DBAs to put different thresholds on different metrics 
tied to a service, and provides alerts dependent on those values.

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.

**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 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 has availability sets: if you put two VMs in one availability set, they will 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 six regions covering North America, Europe, and Asia with plans to introduce South America and Australia in 2017. Each region is represented by multiple zones, allowing you to distribute your apps and storage across to protect against service disruptions. It is 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-premise, 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 M4 instances, which are general-purpose VMs, or on R3 instances, which are memory-optimized VMs. Inside Google Cloud Platform you can choose Cloud SQL using either standard or high-memory machines based on application demands. In Microsoft Azure, database workloads are commonly deployed on DS-Series VMs or GS-Series VMs.

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 Premium SSD storage on three different tiers. 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 is not 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, your business risks becoming locked into a particular cloud provider because of the lack of in-house skills.

This lack of expertise/resources has been cited as the Number One challenge of organizations that are moving to the cloud\(^3\).

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 in 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 database running in the public cloud. As with IaaS, the virtual data center, its physical location, and the physical hardware are managed by the service provider.

Unlike with IaaS, the DBA has nothing to do with the operating system or anything below the operating system. The DBA can scale up or scale down a particular database or instance on demand and can also create or destroy these resources on demand.

The biggest difference with IaaS is that core operational tasks are all provider managed. These core operational tasks include backups, patching, and high availability and disaster recovery.

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

<table>
<thead>
<tr>
<th>DBA role with IaaS</th>
<th>DBA role with DBaaS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manage the OS</td>
<td>No access to the OS</td>
</tr>
<tr>
<td>Install patches</td>
<td>No patching</td>
</tr>
<tr>
<td>Upgrade the system</td>
<td>One-button upgrade</td>
</tr>
<tr>
<td>Perform backup and recovery</td>
<td>Backup and recovery performed by service provider</td>
</tr>
<tr>
<td>Determine and implement HA and DR architectures and strategies</td>
<td>HA and DR are trivial</td>
</tr>
</tbody>
</table>

With all the automation of DBA operational tasks that DBaaS provides, the natural question is, “Why do we still need DBAs?” The answer is that we need DBAs for all the tasks that are very difficult to automate. The new role of DBAs involves code review, performance tuning, custom monitoring, automation and security.
CODE REVIEW
DBAs have a critical role to play in code review. They need to understand the code running inside the database and which pieces are a part of the application’s critical code path. In this way, they can mitigate the risk of a bad piece of code being released to the database. The effects of a bad piece of code released in production can produce a system outage that is not remediated by the general HA mechanisms built by DBaaS providers.

An example is a retailer and submission of orders. The DBA needs to monitor performance specifically for this critical code path at all times. If some bad code is released for the submission of orders, and it’s causing a lot of blocking in the database, the DBaaS monitoring will not detect the problem—and a failover will not solve it either. The DBA needs to review the code, find the bad code, and provide a fix.

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.

As new features are released faster 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 will translate into a permanent bill increase. This cost is often avoidable if the DBA is an experienced database tuner.

DBAs can also provide valuable insights into monitoring requirements that will help keep monitoring costs in line. DBaaS providers offer generic performance-level metrics but don’t have any insight into what is valuable to an organization, from a database perspective, for the application. DBAs can provide value to an organization through this knowledge.
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—and 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 new 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. And these automation templates are just a starting point.

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 in order to respond very 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 an IT service provider, will also be on the lookout for efficiencies.

**HYBRID ARCHITECTURES**

Another area of cloud adoption is the need to build hybrid architectures for both IaaS and DBaaS because most businesses will still keep some components on-premise. According to IDC, more than 70 percent of heavy cloud users are considering a hybrid cloud strategy\(^4\).

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-premise 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-premise, 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 is 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
Databases are important because they are the foundational software of operational and decision support systems. Without the database, any type of analytics or front-end tools simply do not work. Nowadays, company revenue is closely tied to these types of systems—in some cases, it is 100 percent tied to them. For example, if you are an online retailer and you can't receive orders, you are not doing any business if your database and the front end are down.

More generally, for all types of businesses, databases hold the history of the company operations. Any type of business, any type of customer, any type of operation that the company has performed, is stored somewhere in a database. This is very valuable. For many companies, losing this information could mean irreparable damage to their brand and significant loss of business.

Finally, with data warehousing and big data becoming mainstream, operational systems can be turned into actionable information that you can mine and use for competitive advantage. Not doing so will put you at a disadvantage to competitors who may be willing to go that extra mile.

Knowing how important databases are, we see the value of having the best database implementation possible, on-premise or in the cloud. A strong foundation is going to increase the resilience of anything you build on top of it. Your applications, analytics, front end, and eventually the quality of your operations and the decisions you make, are all based on your databases.

Unfortunately, the opposite also applies. A weak foundation at the database level leaves you vulnerable to issues upstream on the other layers of your systems.
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. And 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 very costly incidents. Looking forward, DBAs also need to develop the skills for data-centric development.

THE VALUE OF OUTSOURCING EXPERT DATABASE SERVICES
How does the new, evolving role of DBAs worldwide relate to managed services?

The cloud has a very aggressive release cycle of new features and changes, so speed of execution is essential. In addition, it’s more difficult than ever before to find a single DBA who is an expert in all cloud providers, all database products, and even in all the different facets of a single database product.

The only viable way to face this rate of speed and innovation, and to also continue providing consistent and excellent service, is to use a team-based approach with highly specialized talent—but that can be very expensive.

The role of database managed services is to provide a 24x7, highly available, redundant team of professionals with senior-level skills. These skills have changed and evolved in the cloud, but they are still necessary.

Managed services team members bring added value to clients because of their specialization, professional networks, industry exposure, and strategic advice. The cloud makes this insight and experience more valuable than ever.

So, whether you want to augment your existing DBA staff or offload your entire database management, database managed services may be the most efficient and cost-effective way to proceed. Gartner thinks so: in their list of Top 10 IT cost optimization ideas, Number 3 is “Leverage Cloud Services”[1].

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 are not locked in long-term.

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, 7 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 this 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 are 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.

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 will need to decide whether to augment your existing in-house DBA staff or adopt database managed services—because your mission-critical systems cannot afford any 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.
ABOUT THE AUTHOR

Warner Chaves  
@warchav

Warner Chaves is a principal consultant at Pythian, a Microsoft Certified Master, and a Microsoft MVP. Warner has been recognized by his colleagues for his ability to remain calm and collected under pressure. His transparency and candor enable him to develop meaningful relationships with his clients, where he welcomes the opportunity to be challenged. Originally from Costa Rica, Warner is fluent in English and Spanish.

CONTRIBUTOR

Vladimir Stoyak

Vladimir Stoyak is a principal consultant for big data. Vladimir is a certified Google Cloud Platform Qualified Developer, and Principal Consultant for Pythian’s Big Data team. He has more than 20 years of expertise working in Big Data and machine learning technologies including Hadoop, Kafka, Spark, Flink, Hbase, and Cassandra. Throughout his career in IT, Vladimir has been involved in a number of startups. He was Director of Application Services for Fusepoint, which was recently acquired by CenturyLink. He also founded AlmaLOGIC Solutions Incorporated, an e-Learning analytics company.

ABOUT PYTHIAN

Pythian is a global IT services company that helps businesses become more competitive by using technology to reach their business goals. We design, implement, and manage systems that directly contribute to revenue and business success. Our services deliver increased agility and business velocity through IT transformation, and high system availability and performance through operational excellence. Our highly skilled technical teams work as an integrated extension of our clients’ organizations to deliver continuous transformation and uninterrupted operational excellence using our expertise in databases, cloud, DevOps, big data, advanced analytics, and infrastructure management.