TECH
mahindra

Oracle Database Migration to PostgreSQL on Azure

WHITEPAPER

Abstract

Migrating an Oracle database to PostgreSQL on Microsoft Azure signifies a decisive shift towards a more agile and cost-effective data structure.

This whitepaper is intended for solution architects who want to migrate their Oracle databases to PostgreSQL on Azure cloud environment. The goal is to provide the options of Azure native pathways for PostgreSQL and the alternate pathways with EDB Postgres advanced server on Azure cloud.

Key Takeaways

Introduction

PostgreSQL, is a powerful open-source software (OSS) database system with a large development community. It supports both relational and JOSN data types. Apart from its high compatibility with Oracle, PostgreSQL is being consistently updated by the development community as well as other companies to increase the extensibility for different use cases. It provides users an ever-expanding number of functions that assist technical teams in building applications, maintaining data integrity, and building secure data environments. It can have customized extensions to cater to very specific needs (e.g., extensions to handle time-series, geo-spatial data, and extended logging).

The continuous enhancement of PostgreSQL and adoptability with various platforms make it quite versatile. Most startups and enterprises today use PostgreSQL to power their dynamic websites and applications. And many Business functions across various industries operating with Oracle have been inclined towards PostgreSQL migration over the last decade.

A PostgreSQL database hosted on Azure cloud provides additional advantages of manageability and integration benefits of Azure, enterprise-grade scalability, built-in intelligence, proven resilience, and stability. Besides the migration of Oracle databases to PostgreSQL on Azure cloud can reduce the overall cost of ownership significantly.

The objective of this whitepaper is to provide a direct Azure native pathway for Oracle database migration to PostgreSQL and an alternate pathway of EDB Postgres advanced server on Azure cloud. While the direct Azure native pathway provides a more cost optimized solution, the EDB Postgres advanced server on Azure is an enhanced form with inclusion of many enterprise-class features, improved security controls and is lot more convenient for the Oracle developers and DBAs to adopt.

Major Challenges with Oracle

Oracle being a commercial RDBMS system comes with heavy licensing costs. Additional tools like Golden Gate for real-time replication may need additional licensing. Oracle support too, is not free and must be paid for. It has an integrated authentication system which is not very flexible. While working on Oracle, the developers are restricted in their abilities to seamlessly integrate Oracle components simply by including the header file in the project. Further, the formidable tasks of rolling out updates, upgrades, security patches, monitoring, backups, troubleshoot and manage at scale reduce the efficiency.

In comparison PostgreSQL provides a lot more agile and cost-effective database system. Though both Oracle and PostgreSQL provide application APIs for communicating with databases, the latter's big benefit is that it is open source. A developer can simply include the header files in the project to directly access any of the PostgreSQL components. PostgreSQL is based on host-based authentication, and thus it can support a variety of authentication methods (trust authentication, peer authentication, password authentication). When compared to Oracle, this provides more authentication flexibility and enables delegation of the process.

PostgreSQL Pathways (Deployment models) in Azure Environment

Azure Native PostgreSQL

One can either install PostgreSQL v15.4 or below directly on Azure VM (IaaS) or use Azure database for PostgreSQL (PaaS).

With IaaS option, one can run PostgreSQL Server inside a fully managed virtual machine in Azure cloud environment. This offers full control over the operating system and the PostgreSQL server instance configuration. However, with this the customer will also have to completely manage the VM and many DBA tasks. These tasks include maintaining and patching database servers, database backup and restore, monitoring, and high availability design.

The PaaS model, which is known as Azure database for PostgreSQL, offers a managed database service wherein the users don't get access to the underlying VM and thus don't have to manage it. This service performs automated patching of the underlying hardware, operating system (OS), and database engine. The patching includes security and software updates. For the PostgreSQL engine, minor version upgrades are included in the planned maintenance release. Users can configure the patching schedule to be system managed or define their custom schedule. While provisioning this service, the users can also configure the settings for automated backup, monitoring and High availability as per their requirements.

Both the above service models use the open-source PostgreSQL database engine.

In these native pathways (IaaS or PaaS), PostgreSQL lacks the support for many Oracle-compatible features, e.g., Pro*C, bulk data loader, PL/SQL debugger, and more. Thus, the DBAs and developers working with Oracle for years might find this form of PostgreSQL difficult to adopt. Besides it has limited security features and there is no vendor support available for the database.

Being an open source there are no charges for database. But Azure of course bills for the associated compute and storage elements.

Alternate Option (EDB Postgres Advanced Server) on Azure Cloud

EDB Postgres Advanced Server (EPAS) is built on open-source PostgreSQL, but it is enhanced with many enterprise-class features and tools for backup, disaster recovery, replication, database management, monitoring and tuning. The EPAS platform is designed to meet the needs of large enterprises and to cover more use cases than the open-source community users.

Most of the Oracle-compatible features are supported by EPAS and thus it meets most of the mission-critical requirements for enterprise users. For organizations planning to migrate from Oracle to PostgreSQL, the EPAS compatibility can make the transition much smoother. Besides the Oracle developers and DBAs would find EPAS a lot easier to adopt since it understands and executes Oracle's PL/SQL commands natively without performance issues and without difficult-to-debug emulation.

It has enhanced security features, like row level security, SQL injection protection, and data redaction/dynamic data masking.

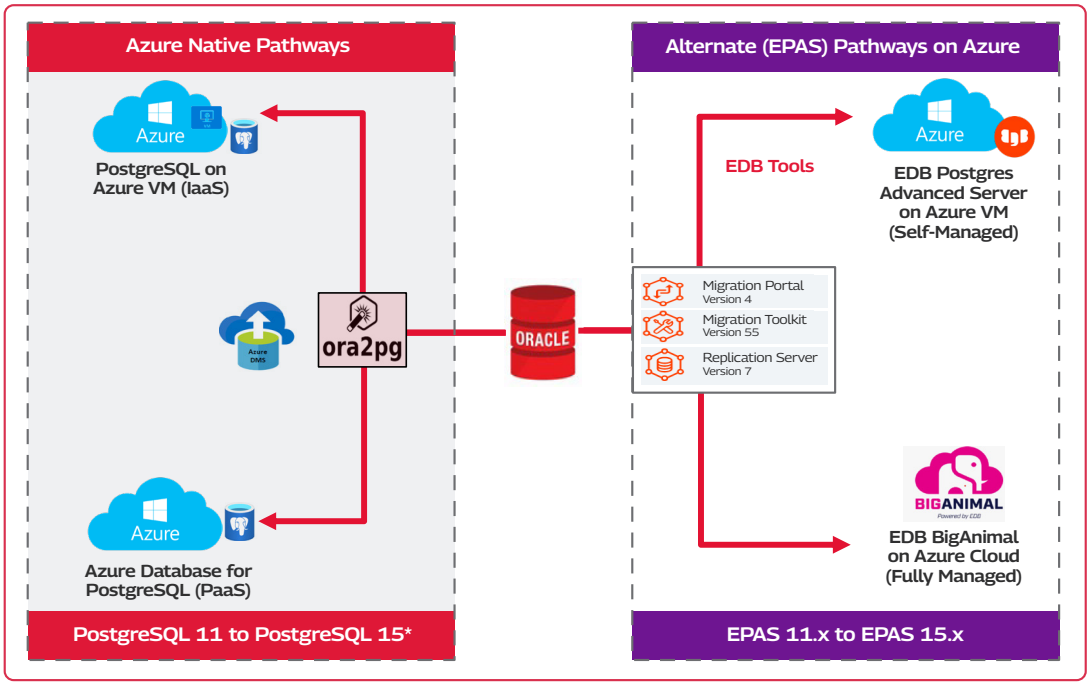
Unlike the open-source PostgreSQL, the EPAS is supported through all phases of the application lifecycle, including application development, testing, production deployment and steady-state operations.

For the billing, the database is charged by EDB based on the number of provisioned vCPUs whereas Azure charges for the infrastructure elements (compute, storage).

EPAS can be deployed with self-managed or fully managed mode in Azure environment.

In the self-managed mode, EPAS (v 11.x to 15.x) is installed in the Azure VM. The customer would get all the benefits of EPAS, but the infrastructure elements will have to be self-managed, Also the operations like backup, monitoring and high availability will have to be configured using EDB tools.

In the fully managed mode, it's known as EDB BigAnimal which provides databases as a managed service without the need to manage any Infrastructure in Azure. The BigAnimal portal and CLI can help deploy the BigAnimal instance in Azure environment. Here the operations like backup and high availability would be available in fully automated form.



Postgre SQL 15 is under Public Priview for Azure database for PostgreSQL Flexible server

Figure 1: Deployment Models for PostgreSQL on Azure

Migration

Azure Native PostgreSQL

Ora2PG (a free tool) is widely used for migration from Oracle to Azure native PostgreSQL.

First an assessment of the on-prem source Oracle environment is done using the Ora2PG to determine the migration complexity and technical effort levels. The outcome report for this assessment should be used by the migration team to determine the practicality of this migration pathway.

During the migration, The Oracle objects are exported as PostgreSQL objects (DDL files) by running the relevant Ora2PG script and imported to the target PostgreSQL environment. Similarly, the data is exported and imported to PostgreSQL using the appropriate Ora2PG command.

During this migration process, there are high chances of some of the objects (e.g., stored functions, triggers, and views.) being unable to be identified properly. These objects would require a significant amount of code rework in the target environment.

Finally, the data is sync is done to mark the migration completion and start cut over. For the data sync of each table, query is used to filter (cut) data by date, time, or another parameter. Then a second query is used to migrate the remaining (delta) data. This process being manual is relatively cumbersome.

EPAS on Azure Cloud

The migration process is more convenient with EPAS when compared to Ora2PG for Azure native PostgreSQL.

EDB Migration portal is used for migration of data objects (tables, views etc.) and code objects (functions, stored procs, packages) and EDB migration toolkit (MTK) can be used for data migration. Many of the incompatible objects are auto repaired by tool and migrated properly and the count of invalid objects in the target environment is considerably less. Thus, relatively much less code rework is required for these objects.

The EDB replication server (which is used in conjunction with EPAS) supports change data capture and thus it can be used for a much easier and automated data sync (through replication) after the initial full migration. This is a significant advantage over Ora2PG for Azure native PostgreSQL, where the data sync is performed manually.

Decision-making Process

The decision to choose between Azure native PostgreSQL or EPAS on Azure cloud should be based on multiple considerations. Here are some of the important factors to help choose between the two.

Is Oracle compatibility an important criterion for the Developers?

For Oracle developers and DBAs, the biggest differentiator with EPAS is that it

understands and executes Oracle's PL/SQL commands natively without performance issues and without difficult-to-debug emulation. Hence the Oracle-skilled resources hardly find EPAS difficult to adopt. Besides, EPAS supports most of the other Oracle compatible features. In comparison, the Azure native PostgreSQL (IaaS or PaaS) lacks the support for many Oracle-compatible features, e.g., Pro*C, bulk data loader, PL/SQL debugger, and more. Thus, the Oracle developers and DBAs find this form of PostgreSQL difficult to adopt.

Do we need Enterprise tools for Backups, Monitoring, Replication, HA etc.?

When compared with Azure native PostgreSQL, EPAS offers a more robust set of Enterprise tools for backup, high availability, DR, replication, database management, monitoring and tuning. This puts EPAS in the pole position to meet the needs of large enterprises and to cover more use cases.

As per Ora2pg assessment report is the Migration Effort level B or C AND Technical complexity level 4 or 5?

Ora2pg assessment tool is commonly used for assessing the source Oracle environment before migration to PostgreSQL. It gives an output in the form of a report which summarizes technical complexity of the source database and the efforts required for rewriting the codes of invalid/unidentified objects in PostgreSQL database. The efforts are represented on a scale of A to C and the database technical complexity level is denoted on a scale of 1 to 5. If the efforts emerge to be B or above and technical complexity level transpire to be 4 or above, it will mean significant amount of work is needed in the PostgreSQL database. In such scenarios, EPAS on Azure cloud could be a better choice since with EPAS, many of the incompatible objects are auto repaired by the EDB tools and migrated properly. So, the number of invalid objects in that target environment is considerably reduced and hence relatively much lesser code rework is required for these objects.

Do we require 24*7 support for the database?

The EPAS on Azure gets the vendor support through all phases of the Application lifecycle, including application development, testing, production deployment and steady state operations. On the other hand, Azure native PostgreSQL must be self-supported since it doesn't get any vendor support for the database.

If the answer to any of these questions is YES, then EPAS on Azure is a more appropriate selection. But if most of the answers are NO, then Azure native PostgreSQL is a better choice.

Conclusion

As more and more organizations explore Oracle database migration to Azure environment, PostgreSQL certainly can be a good alternative offering more agile, flexible, and cost-effective data structure in Azure cloud. However, it is very important to have a detailed assessment of the source Oracle environment to finalize the right target (Azure native PostgreSQL or EPAS on Azure Cloud). Azure native pathway provides a more cost optimized solution. But with EDB Postgres advanced server, one gets many Enterprise-class features and improved security controls. It is also important to choose the right service model with either pathway, i.e. IaaS or PaaS (with Azure native PostgreSQL) and Self-managed or fully managed (with EPAS on Azure).

Tech Mahindra can help customers in determining the appropriate Target in Azure cloud by a thorough assessment of the source oracle environment, capex and opex roadmap for upcoming few years, security & compliance principles followed at the respective business unit and experience/skill sets of team members. Tech Mahindra can also assist in the actual migration process to Azure and can provide managed services for database management. The support with managed services can have three categories, i.e., Level 1, Level 2 and Level 3 based on the complexity of the database and nature the support required.

Authors

Guru Prasad C P

Group Practice Head, Azure Cloud Services

Guru Prasad C P has an experience of over 22 years with over 8 years specifically in the public cloud working in Asia, ANZ, Europe, and the US. His experience includes, setting up practice teams aligned to industry verticals and horizontals, analyst interactions for positioning the offerings, hiring the right talent, involving in strategic exercise mergers and acquisitions, organization building, creating frameworks & IP's. At Tech Mahindra he is responsible for practice and competency development which includes alignment with OEMs for solutions, offerings and adoption of new technologies, customer interfacing where he acts as a trusted advisor in providing unbiased views/opinions and aligning with organization goals at the same time, value creation, developing practice areas deal making, solution support for large deals, and carve out deals from azure and hybrid cloud perspective Guru has a keen interest and expertise in verticals including travel, transport, manufacturing, insurance, educational and government charitable trusts.

Sateesh Mohapatra

Technical Architect (Azure Cloud COE)

Sateesh Mohapatra is an enterprise architect with over 22 years of IT experience predominantly in public/hybrid cloud environment. He has worked on complex projects involving discovery, assessment, foundation build and migrations of applications (with optimization/modernization) to public cloud environment. He has also gained rich experience on pre-sales activities involving estimations, solutioning and proposals for public cloud migration projects. Besides he has worked as an Oracle DBA for 9+ years with hands-on experience on very large and complex environments.

Naveen Jogappa

Solution Architect (Azure Cloud Services))

Naveen Jogappa is an Azure cloud solution architect with 16 years of experience in public, private and hybrid cloud solutions. He is experienced in enterprise cloud strategies, pre-sales solutions, and technical architecture skills. At Tech Mahindra he is part of the Azure cloud practice and responsible for competitive solutions across cloud migration, modernizations, and hybrid solutions.