As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology
Volume 9, Issue 08, pages 35-40
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol9issue09.html
DOI: https://doi.org/10.46501/IJMTST0909007

Database Migration Services in Cloud Services

Bairysetti Prasad Babu[1] | Dr. Prasadu Peddi[2] | Dr. Suneel pappala[3]

[1]Research Scholar, Dept of Computer science & Engineering, Shri JJT University-Rajasthan.
[2]Dept of Computer science & Engineering, Shri JJT University- Rajasthan
[3]Associate Professor. Dept of Computer science & Engineering, Lords Institute of Engineering And Technology, Hyderabad.

To Cite this Article
Bairysetti Prasad Babu, Dr. Prasadu Peddi and Dr. Suneel pappala. Database Migration Services in Cloud Services, 2023, 9(09), pages. 35-40. https://doi.org/10.46501/IJMTST0909007

Article Info
Received: 06 August 2023; Accepted: 24 August 2023; Published: 10 September 2023.

ABSTRACT

A key challenge in porting enterprise software systems to the cloud is the migration of their database. Choosing a cloud provider and service option (e.g., a database-as-a-service or a manually configured set of virtual machines) typically requires the estimation of the cost and migration duration for each considered option. Many organizations also require this information for budgeting and planning purposes. Existing cloud migration research focuses on the software components, and therefore does not address this need. We introduce a two-stage approach which accurately estimates the migration cost, migration duration and cloud running costs of relational databases. The first stage of our approach obtains workload and structure models of the database to be migrated from database logs and the database schema. The second stage performs a discrete-event simulation using these models to obtain the cost and duration estimates. We implemented software tools that automate both stages of our approach. An extensive evaluation compares the estimates from our approach against results from real-world cloud database migrations.

Keywords: Database, Cloud migration, Cloud Database Migrations, Software components, Cloud Deployment.

1. INTRODUCTION

The benefits of hosting an enterprise system on the cloud — instead of on-premise physical servers — are well understood and documented. Some organizations have been using clouds for over a decade and are considering switching provider, while others are planning an initial migration. In either case, the most challenging component to migrate is often the database due to the size and importance of the data it contains. However, the existing cloud migration work focuses on the software components and gives minimal consideration to data. For instance, the ARTIST and REMICS cloud migration methodologies refer to the database but do not support any database specific challenges. Similarly, cloud deployment simulators like CDO Sim focus only on compute resources. The limitations of these existing cloud migration methodologies are described further in "Related work" section.

Migrating large relational databases from physical infrastructure into the cloud presents many significant challenges, e.g., managing system downtime, choosing suitable cloud instances, and choosing a cloud provider. The database could be deployed on a database-as-a-service offered by one of several public cloud providers, or installed and configured on a virtual machine(s). With either option, selecting the appropriate cloud resources requires knowledge of the database workload and size. The infrastructure of the source database may impact the migration duration; if it has limited available capacity or bandwidth, then it will take longer to extract the data. An organization may wish to upgrade the existing database hardware to speed up migration, or schedule downtime to migrate the database while it is idle.

In this work, we assist with this decision-making process via a tool-supported approach for evaluating cloud database migration options. Our approach has two stages—database workload and structure modelling, and database migration simulation—and estimates migration duration, migration costs, and future cloud running costs.We assume the source and target databases have an identical: schema, type (e.g., relational or NoSQL), vendor (e.g., Oracle or MySQL), and software version. Changing any of these parameters is a complex activity, which organizations tend to perform separately (as discussed in "Approach overview" section).

2. LITERATURE REVIEW

Abou_el_ela Abdou Hussein (2021) Data Migration is a multi-step process that begins with analyzing old data and culminates in data uploading and reconciliation in new applications. With the rapid growth of data, organizations constantly need to migrate data. Data migration can be a complex process as testing must be done to ensure data quality. Migration also can be very costly if best practices are not followed and hidden costs are not identified in the early stage. On the other hand, many organizations today instead of buying IT equipment (hardware and/or software) and managing it themselves, they prefer to buy services from IT service providers. The number of service providers is increasing dramatically and the cloud is becoming the preferred tool for more cloud storage services.

Abel Adane (2020) Several researches and case studies have been revealing a repetitive issue and that is monopolization of controlling, measuring, and metering of the consumer's billing for the consumed Cloud Services. The Cloud Service Providers (CSPs) have been hiding, or monopolizing the controlling, metering tools and ignoring the calculation and consideration of the service defects like downtime, service outage, poor performance and client service migration delays during overloads of the servers. The main issues behind these discrepancies are the Service Level Agreement (SLA), its violation, and the monopolized Cloud Governance. Till now there is a lack of parallel monitoring and metering system of the consumed cloud services at the customer level under the provisions of SLAs.

Shuying Zhai (2020) The innovation of this paper is to propose a new impact path that by attracting migrant population, basic public services can stimulate national economic growth significantly. Based on the "cost-benefit" theory, this paper takes basic public services into the population migration decision-making model, then builds an endogenous economic growth model by using population immigration as an increase of human capital. Theoretical analysis finds that basic public services can attract people to move in positively and promote regional economic growth indirectly. Based on the above conclusions, this study presents that, we should increase fiscal transfer payments in backward areas, develop and improve the equalization system of basic public services while adapting to local conditions.

Arif Iqbal (2019) Cloud data migration is the procedure of moving information, local host applications, services, and data to the distributed cloud computing infrastructure. The success of this data migration process is depending on several aspects like planning and impact analysis of existing enterprise systems. One of the most common operations is moving locally stored data in a public cloud computing environment. This paper, through a multivocal literature review, identifies the key advantages and consequences of migrating data into the cloud. The results of this research paper can give a road map for the data migration journey and can help decision makers towards a safe and productive migration to a cloud computing environment.

Database Modelling Stage

Systems being migrated to the cloud (and their databases) will usually have been in continuous development for a long time. This can result in unused tables or data, which engineers are reluctant to clean-up for fear of unintended consequences. This issue is compounded when there is poor documentation or when knowledgeable engineers leave the team. However, an accurate understanding of the database must be obtained to choose cloud migration parameters and approaches. This information is also required for other common activities, such as database refactoring, archival of old data, and potential transition to a NoSQL datastore.

A key challenge when developing a database modelling method is heterogeneity, as the available tools and features that the new method can exploit are different between each database provider. For example, MySQL Enterprise Monitor can generate statistics on the workload and MySQL Workbench can model the structure. Other databases have similar tools with varying functionality. As described next, our database modelling method and DBLModeller tool adopt a platform-independent approach that utilises a SQL schema dump and a SQL query log (if necessary). As we show later in the evaluation section, this approach overcomes many of the differences between SQL dialects.

Migration Simulation Stage

The second stage of our approach uses discrete-event simulation to estimate the database migration cost, migration duration, and future cloud running costs. Our simulation method, called MigSim, focuses exclusively on the database and considers: (1) its size and workload, (2) growth trends, and (3) compute instance/virtual machine performance. Accurate estimates of the database migration costs enable organizations to plan, budget and investigate trade-offs. From a planning perspective, the organization will typically want to know how much time a migration requires, as this may rule out an Internet-based migration. One common trade-off an organization might want to investigate is duration versus cost. Additional bandwidth or increased database performance could speed-up the data transfer into the new database. Similarly, they can look at the cost benefits of 'cleaning-up' the database before migration, i.e., identifying and removing unneeded tables or archiving old data. Our simulation method can be equally applied to migrations between clouds, or from on-premise databases to the cloud.

Cloud cost Metamodel

We developed a metamodel to define the structure and content of our cloud cost model. The wide ranges of services typically offered by cloud providers and the regional variations in pricing are complex; this makes a model-based approach for including cost data in the simulation desirable.

Our Metamodel — implemented in the Eclipse Modelling Framework — has been designed to support charges from major public clouds which may be incurred during database migration. Each instance of our metamodel corresponds to a single cloud provider and covers a set of cloud services (e.g., Amazon RDS and EC2). Finally, each cloud service is associated with multiple cloud charges.

Completeness, correctness, and performance

Model completeness, model correctness, and model extraction performance have been evaluated together due to their interdependence (e.g., extracting a complete model will require more time than an incomplete model). DBLModeller was compared to Gra2MoL's PLSQL2KDM example as this had the highest level of SQL support at the time of writing. We extracted KDM models from the database schemas of four systems: Apache OFBiz, MediaWiki, Science Warehouse, and a student record system. With OFBiz and MediaWiki we obtained Oracle and MySQL versions of the schema by installing them on both databases.

Using Arcion for Migration

Aside from using the cloud provider tools for data migration, there is also an easy way to migrate data by using Arcion. Arcion supports a wide number of source and target DB engines and big data platforms. Using Arcion, you can migrate data to your new platform while enjoying the benefits of zero-downtime and automatic schema migration. This allows you to pick a target platform, NoSQL or SQL, and easily migrate the data and the schema automatically. Another benefit is zero-downtime and zero data loss since you can also enable data replication. With data replication enabled, you can migrate your data and even keep the latest data that changed after the migration started and have it kept up to date on the new database. When you decide to flip application and services to use your new database, you can be sure that even the latest data is available.

3. RESEARCH METHODOLOGY

In this study, we have considered a Private Cloud, and two Public Clouds namely Amazon Web Services (AWS), and Google Cloud Platform (GCP). The user is initially authenticated in two public cloud providers. Create an instance of the chosen services after the account is created and authenticated. For future cloud access, the service account's login credentials and private key are downloaded. The required softwares are installed on the cloud instance which is necessary for the analysis after authenticating the user on both public clouds. Using MySQL Workbench, create a database and link it to the cloud. The Migration Manager receives a migration request from the user and the manager then connects with Cloud Service Provider to gain database access and establish a connection. Once the connection is secure, the backup of data is created and encrypted by the Inter Cloud Framework (ICF). The Migration Manager establishes a secure connection with the destination cloud to migrate the file from the source cloud provider to the destination cloud provider. Upon sending the files the destination cloud provider decrypts, inflates and restores the data on the Amazon Relational Database Service (RDS) service. The demonstration of live migration of Applications and Databases between the Clouds Amazon Web Services (AWS) and Google Cloud Platform (GCP) is accomplished. During the application migration and database migration, CPU and RAM utilisation readings are recorded and used to analyze the system with the variation of database and application sizes.

4. RESULTS

Table 1 tabulates the file size and virtual memory size taken into consideration for migrating the VM.

Table 1: Modelling of file size vs. virtual memory size

	File Size (MB)	VM Size (Byte)
1	100 KB	261577984
2	200 KB	274261760
3	300 KB	276341888
4	400 KB	277746240
5	500 KB	278639744
6	600 KB	279588864
7	700 KB	280581824
8	800 KB	281691840
9	900 KB	281941120
10	1000 KB	282129536

The deviance residual of the virtual memory size is discussed in Table 2.

The coefficients of virtual memory are discussed size in Table 3.

Table 2: Deviance residuals of virtual memory size

Min	1Q	Median	3Q	Max
-220.815	-146.051	-6.776	145.556	260.397

Table 3: Coefficients of virtual memory size

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	-1.069e+04	2.758e+03	-3.877	0.00469
Committed Virtual Memory Size	4.052e-05	9.937e-06	4.077	0.00355

The residual standard error is 183 on 8 degrees of freedom. Multiple R-squared has a value of 0.6751 with an Adjusted R-squared of 0.6345. F-statistics observed is 16.62 on 1 and 8 degrees of freedom and the p-value is 0.003547.

Table 4 tabulates the file size vs CPU load of the process taken for migration in bytes. As seen in the result tabulated the framework maintains minimum load making CPU available for other processes.

Table 4: Modelling of file size vs. process CPU load

	File size	Process CPU load in bytes
1	100 KB	0
2	200 KB	0.241982401
3	300 KB	0.242626359
4	400 KB	0.242714992
5	500 KB	0.249651686
6	600 KB	0.249876861
7	700 KB	0.25000672
8	800 KB	0.250076824
9	900 KB	0.254459951
10	1000 KB	0.264459951

The x-axis represents the file size in Kilo Bytes and the y-axis represents the Process CPU Load in bytes. The file size is varied from 100KB to 1000KB and the Process CPU load is evaluated. It is observed that with the increase in file size, the CPU load is almost constant with less.

Table 5 shows the deviance residuals of the process CPU load.

Table 5: Deviance residuals of process CPU load

Min	1Q	Median	3Q	Max
-388.94	-169.46	23.05	167.98	360.74

Table 6 shows the coefficients of the process CPU load. It is evident from the table that the model is 86% effective thereby reducing the CPU load.

Table 6: Coefficients of process CPU load

	Estimate	Std. Error	t-value	Pr(>|t|)
(Intercept)	47.28	259.50	0.182	0.8600
Process CPU Load	2238.44	1095.80	2.043	0.0754

The residual standard error is 260.3 on 8 degrees of freedom. The multiple R-squared obtained is 0.3428 with an Adjusted R-squared value of 0.2606. F-statistics is 4.173 on 1 on 8 degrees of freedom, and the p-value is 0.007535.

For File Sizes varying from 100 KB to 1000KB, the CPU time is calculated. File Size vs CPU Time is tabulated in Table 7. The file size is measured in Kilo Bytes whereas Process CPU time is measured in terms of nanoseconds.

Table 7: Modelling of file size vs. process CPU time

	File size	Process CPU time in nanosec
1	100 KB	14468750000
2	200 KB	14500000000
3	300 KB	14562500000
4	400 KB	15000000000
5	500 KB	15593750000
6	600 KB	15890625000
7	700 KB	16375000000
8	800 KB	16562500000
9	900 KB	16625000000
10	1000 KB	16703125000

Table 7 shows a gradual increase in time with the increase in file size. However, the model shows the CPU time as a gradual slope rather than a spike with an increase in file size with the Process CPU time maintained at the lowest.

5. CONCLUSION

Cloud computing ensures the on-demand availability of the system resources via ubiquitous network access enabling a multi-tenant model where capabilities are elastically provisioned cost-effectively. The migration of data and applications among cloud service providers and consumers has given rise to newer challenges to ensure interoperability and portability. Cloud users still find that the live migration of VMs is still challenging. The demonstration of live migration of applications and data between two public clouds using the proposed Inter Cloud Framework is discussion. Any application and data on the Cloud are difficult to migrate between two different public clouds. The Inter Cloud Framework helps Cloud users migrate between two public clouds with ease and explore migration capabilities between two different public clouds. The experimental setup was made to transfer an application and data between Amazon Web Service and Google Cloud Platform. The proposed framework helps gather the required information and data between the source and target cloud, enabling a smooth migration between them by reducing the usage of system resources. The Inter Cloud Framework helps cloud users migrate with minimum service disruption across many public cloud service providers. In this approach, live migration is demonstrated, and evaluation is done on the Amazon S3 instance.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Richard F. Paige (2018),"Evaluating cloud database migration options using workload models",Journal of Cloud Computing,ISSNno:2192113X, Vol.7,https://doi.org/10.1186/s13677-018-0108-5

[2] Arif Iqbal (2019),"Key Opportunities and Challenges of Data Migration in Cloud: Results from a Multivocal Literature Review",Procedia Computer Science,ISSNno:1877-0509,Vol.164,Pages.45-55.https://doi.org/10.1016/j.pro cs.2019.12.153

[3] Abel Adane (2020),"Cloud Services Quality Enhancement through Control & Monitoring of Service Level Agreements at Server and Consumer Side", Journal of Computer and

Communications,ISSNno:2327-5227,Vol.8,No.8, Pages.78-90.

[4] Shuying Zhai (2020),"Basic Public Services, Population Migration and Regional Economic Growth",Modern Economy,ISSNno:2152-7261,Vol.11, No.4,Pages.777-784.

[5] Abou_el_ela Abdou Hussein (2021),"Data Migration Need, Strategy, Challenges, Methodology, Categories, Risks, Uses with Cloud Computing, and Improvements in Its Using with Cloud Using Suggested Proposed Model (DMig 1)",Journal of Information Security,ISSNno:2153-1242,Vol.12,No.1, Pages.79-103.

[6] Chiang Lee (2016),"An Adaptive Video-on-Demand Framework for Multimedia Cross-Platform Cloud Services",Journal of Software Engineering and Applications,ISSNno:1945-3124,Vol.9,No.5,Pages.155-174.

[7] Adela Chávez (2014),"Myths and Realities about Mexico-United States Migration",iBusiness,ISSNno:2150-4083,Vol.6,No.2,Pages.63-70.

[8] Fatimah Khalil Aljwari (2022),"Challenges of Privacy in Cloud Computing", Journal of Computer and Communications,ISSNno:2327-5227,Vol.10,No.12, Pages.51-61.

[9] Irfan Ali Jamali (2016),"Data Security of Mobile Cloud Computing on Cloud Server",Open Access Library Journal,ISSNno:2333-9721,Vol.3,No.4,Pages.1-11.

[10] Evelyn Amankwah (2017),"Evaluation of Public Weather Services by Users in the Formal Services Sector in Accra, Ghana",Modern Economy,ISSNno:2152-7261,Vol.8,No.7,Pages.921-945.