ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

# A REVIEW ON DATA MIGRATION IN PUBLIC CLOUD

**Gurdeep Singh*, Dr.Vijay Laxmi**
* UCCA Guru Kashi University Talwandi sabo,India
UCCA Guru Kashi Uneversity Talwandi Sabo, India

**DOI**: 10.5281/zenodo.546319

## ABSTRACT

Cloud computing has recently became a widely discussed topic in the IT industry. More and more organizations consider using the Cloud, because it enables an easy and cost efficient way of hosting applications, with dynamic scaling and geographical distribution possibilities. Still, it is not clear how and when cloud computing should be used. Existing application are often written in a way that does not really fit a cloud environment well. Also, certain quality attributes (e.g. performance, security or portability) can be affected. More studies are needed on how existing systems should be plugged into the Cloud and what are the consequences of the migration. Data migration and application migration are one of popular technologies that enable computing and data storage management to be autonomic and self-managing. we examine important issues in designing and developing scalable architectures and techniques for efficient and effective data migration and application migration. The first contribution we have made is to investigate the opportunity of automated data migration across multi-tier storage systems.

**KEYWORDS**: Cloud computin; public cloud platform; Data migration, Enterprise application component.

## INTRODUCTION

Computing and communication have continued to impact on the way we run business, the way we learn, and the way we live. The rapid technology evolution of computing has also expedited the growth of digital data, the workload of services, and the complexity of applications. Today, the cost of managing storage hardware ranges from two to ten times the acquisition cost of the storage hardware. We see an increasing demand on technologies for transferring management burden from humans to software. Data migration and application migration are one of popular technologies that enable computing and data storage management to be autonomic and self-managing.

A unique feature of our automated migration approach is its ability to dynamically adapt the data migration schedule to achieve the optimal migration effectiveness by taking into account of application specific characteristics and I/O profiles as well as workload deadlines. Our experiments running over the real system trace show that the basic look-ahead data migration model is effective in improving system resource utilization and the adaptive look-ahead migration model is more efficient for continuously improving and tuning of the performance and scalability of multi-tier storage systems. The second main contribution we have made in this dissertation research is to address the challenge of ensuring reliability and balancing loads across a network of computing nodes, managed in a decentralized service computing system. Considering providing location based services for geographically distributed mobile users, the continuous and massive service request workloads pose significant technical challenges for the system to guarantee scalable and reliable service provision. We design and develop a decentralized service computing architecture, called Reliable GeoGrid, with two unique features. First, we develop a distributed workload migration scheme with controlled replication, which utilizes a shortcut-based optimization to increase the resilience of the system against various node failures and network partition failures. Second, we devise a dynamic load balancing technique to scale the system in anticipation of unexpected workload changes. Our experimental results show that the Reliable GeoGrid architecture is highly scalable under changing service workloads with moving hotspots and highly reliable in the presence of massive node failures. The third research thrust in this dissertation research is focused on studying the process of migrating applications from local physical data centers to Cloud. We design migration experiments and study the error types and further build the error model. Based on the analysis and observations in migration experiments, we propose the CloudMig system which provides both configuration validation and installation automation which effectively reduces the configuration errors and installation complexity.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

## DATA MIGRATION IN CLOUD COMPUTING

The significant IO improvements of Solid State Disks (SSD) over traditional rotational hard disks makes it an attractive approach to integrate SSDs in tiered storage systems for performance enhancement. However, to integrate SSD into multitiered storage system effectively, automated data migration between SSD and HDD plays a critical role. In many real world application scenarios like banking and supermarket environments, workload and IO profile present interesting characteristics and also bear the constraint of workload deadline. How to fully release the power of data migration while guaranteeing the migration deadline is critical to maximizing the performance of SSD enabled multi-tiered storage system. In order to fully capitalize on the benefits of SSDs in a multi-tiered storage system with SSDs working as the fastest tier, it is important to identify the right subset of data that needs to be placed on this tier given the limited capacity of SSD tier due to high cost per gigabyte. Specifically, we want to maximize overall system performance by placing critical, IOPS (input/output operations per second) intensive and latency-sensitive data on the fast SSD tier through two-way automated data migration between SSDs and HDDs. By working with a variety of enterprise class storage applications, we observe that many block-level IO workloads exhibit certain time-dependent regularity in terms of access patterns and temperature of extents (hot or cold). For example, in banking applications, IO workloads for account access and credit verification are typically heavier during certain hours of a day. However, such patterns may change from day-time to night-time, from day to day, from weekdays to weekends or from working days to public holidays. Thus, block-level IO profiling is the first step for building an automated data migration system. The next big challenge is to devise strategies

## LITERATURE SURVEY

Mr. Shrikant D. Bhopale [1], Cloud computing is one of the emerging fields in the computer world these days. Cloud computing is attracting everyone with its benefits. Now companies are shifting their focus onto cloud computing. But to be a part of cloud computing environment and to take advantages of cloud computing, legacy applications need to be migrated to cloud. Cloud migration is the process of transitioning all or part of a company's data, applications and services from onsite computers behind the firewall to the cloud or moving them from one cloud environment to another. After migrating to the cloud, the information will be available on the internet so that more people can have access to it as needed.

N. Kishore[2], Everything in Cloud is an emerging concept, with work in multiple areas right from the Infrastructure, to middleware, to applications and to data security. Every domain has its own flavor of Cloud. Technologies and techniques for Cloud has come a long way and is still evolving. Organizations are fast recognizing the value of migrating to Cloud but the biggest concern is around data security. The migration of sensitive data to a public cloud domain has risks associated with data loss, information theft, confidentiality and other vulnerabilities. There are security measures deployed at multiple points but the question of secured end to end data transmission still remains unanswered. Objective is to understand the security parameters that can help build a framework for secured data transmission. With reference to many research papers and reviews, the analysis is to focus on the security of data in transit. Pragmatic observation is that the data at rest in the Enterprise and the Cloud have fairly good security measures, but the data in transit is vulnerable. There is a future scope to build a security framework for the transit data. This paper will highlight few of the security aspects of Cloud that have been developed and the gap around the security of data being migrated from Enterprise storage to Cloud Storage.

Rabi Prasad Padhy[3], Cloud computing is an architecture for providing computing service via the internet on demand and pay per use access to a pool of shared resources namely networks, storage, servers, services and applications, without physically acquiring them. So it saves managing cost and time for organizations. Many industries, such as banking, healthcare and education are moving towards the cloud due to the efficiency of services provided by the pay-per-use pattern based on the resources such as processing power used, transactions carried out, bandwidth consumed, data transferred, or storage space occupied etc. Cloud computing is a completely internet dependent technology where client data is stored and maintain in the data center of a cloud provider like Google, Amazon, Salesforce.som and Microsoft etc. Limited control over the data may incur various security issues and threats which include data leakage, insecure interface, sharing of resources, data availability and inside attacks. There are various research challenges also there for adopting cloud computing such as well managed service level agreement (SLA), privacy, interoperability and reliability. This research paper outlines what cloud computing is, the various cloud models and the main security risks and issues that are currently present within the cloud computing industry. This research paper also analyzes the key research and challenges that presents in cloud

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

computing and offers best practices to service providers as well as enterprises hoping to leverage cloud service to improve their bottom line in this severe economic climate.

S. Dewan[4],Cloud Computing is evolving as future generation architecture for computing. Usually cloud computing is defined as a grouping of computing recourses accessible via internet services. Traditionally the client/user store data in data centres with firewall and other security techniques used to protect data against intrudes to access the data. The client/user has less control over the stored data on the cloud, since the data is stored anywhere across the globe. It is the responsibility of cloud service providers to protect the client/user data from unauthorized access & disclosure and be faithful for the development of cloud computing. A third party auditor (TPA) can be used to provide security services over the cloud and ensures client/user that his/her data is secure from unwanted and suspicious activity. The TPA would not store any kind of data at its end, and it is only confined to providing security service between cloud service provider and the user. During migration of data from one cloud to other cloud, the application framework will provide data integrity verification by using hashing algorithm like SHA-1 and provide encryption/decryption using symmetric algorithm like AES. We have proposed a secure mechanism to secure data migration between clouds using third party auditor (TPA).The proposed approach adapts the secure socket layer (SSL) protocol to increase the level of trust and security for both user and cloud service provider. Network simulation tool (NS-2) is used to analyze the fairness of the channel during data migration between clouds in terms of block encryption/decryption time duration. Our proposed mechanism improves the time constraint parameter for migrating data from one cloud to other.

## DATA MIGRATION SERVICE MODELS

**PaaS Migration**

To use a PaaS cloud service as the target for migration, the application itself must be designed for one or more runtime environments available in the target PaaS service. An example of such an application is one where the business logic is implemented as a set of components which run on an application server (such as IBM's WebSphere, Oracle's WebLogic, or the JBoss server) in combination with a database (such as the Oracle database or IBM's DB2) containing the application's data and also possibly associated code in the form of stored procedures.

In general, a PaaS solution must provide the elements of the particular software stack required by applications such as the operating system, an application server and a database, so that the customer only has to be concerned with the specific application components and data. One must also ensure that the PaaS environment offers the configuration(s) required by the application. This may include software levels, the ability to run scripts, and the presence of certain tools for setup, reporting, monitoring, etc., identical or similar to those present before migration.

**IaaS Migration**

To migrate an application to an IaaS service, the requirements on the cloud service itself tend to be lower. The entire software stack is migrated: the application code itself, plus any supporting code it requires – including the underlying operating system. To achieve this, it must be possible to package the complete software stack as one or more virtual machine (VM) images, which can then be copied into the cloud service and executed there.
Whether the software stack involved will work in a virtual machine environment may depend on whether there is use of specialized device drivers or hardware devices that are unlikely to be supported by an IaaS provider; an application depending on these capabilities is not a good candidate for migration. This can be tested by preparing the virtual machine image of the application and of the supporting code and attempting to execute it on a trial VM environment (either in-house or a test system offered by the cloud service provider). Hidden dependencies can thus be found, corrected, and the process repeated until successful or until it is determined that there is no affordable solution.

**Data Migration techniques in public clouds**

The amount of time it takes to complete the actual migration of objects and data from one database is relatively less than the amount of time it takes to complete an overall migration from assessment to production rollout. Migrations of one relational database to another are comparatively easier than migrations of a non- relational database to a relational database, because the organization of objects in a relational database is quite similar compared to non-relational databases such as hierarchical and network databases. All major relational database vendors also offer tools that provide robust migration capabilities in an automated fashion. Regardless of the level of automation and success factor of any migration tool, however, sometimes manual intervention will be required

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

when migrating from one database to another. Database migration tasks can be divided into the following categories:

- Database schema migration
- Data migration
- Database stored program migration
- Application migration
- Database administration script migration.

Of all the migration tasks listed, the application migration task requires the most manual effort, although new tools and technologies are being developed to facilitate this task.

**Database Schema Migration**

Database schema migration essentially involves migration tables, indexes, and views in a database. Relational databases are similar in terms of how their data is organized in tables and indexes, but they are different in terms of additional extensions to these tables and indexes that are designed to improve performance and facilitate development. Most migration tools can convert the database schema relatively quickly and accurately. Target-specific database schemas can also be generated from modeling tools such as Erwin.

**Data Migration**

Data Migration After database schema migration, some representative data from the source database is migrated to the target database to enable testing and to ensure that the data migration scripts or tools chosen for the task are configured properly. The most common approach for data migration is undoubtedly the use of scripts that execute database utilities to export data from the source database and import it into the target database (Oracle), because they are easy to use and are free. Regardless of the tools and scripts used to perform data migration, migrations of very large databases require planning. When migrating very large databases (those with at least a few terabytes of data) it is important to have the right data migration strategy, have the appropriate tools, and, most importantly, use appropriate database features such as partitioning and compression. Migration of large databases is fraught with challenges, among them a narrow window of time and lack of system resources (e.g., staging areas for data files). The following data extraction and loading strategies can optimize the data extraction, transfer, and loading processes:

- Parallel extraction of data from the source database
- Loading of data into the target database in parallel
- Using multithreaded processes for data loading
- Avoidance of index maintenance during the data loading process
- Reduction of I/O operations and use of staging areas via named pipes for data transfer between source and target databases

**Database Stored Program Migration**

The task of migrating database stored programs includes migration of stored procedures, triggers, and views which, in many relational databases, are used for implementing critical business logic. In databases such as Microsoft SQL Server and Sybase, stored procedures and triggers are used extensively by developers to support simple functions (e.g., the CRUD operations CREATE, READ, UPDATE, and DELETE). However, using stored procedures exclusively for CRUD operations can result in inflexibility because the type of operation executed against a table is limited by the functionality implemented in the stored procedure.

**Application Migration**

Application migration or porting can result from either migrating an application from one environment to another due to a complete rewrite, or simply from an underlying database platform that is being migrated to a new platform such as Oracle.

## CONCLUSION

In this paper, we have discussed various techniques for Data Migration in Public Cloud. Cloud migration is the process of moving data, applications or other business elements from an organization's onsite computers to the cloud, or moving them from one cloud environment to another.

Cloud migration sometimes involves moving data or other business elements between cloud environments, which is known as cloud-to-cloud migration. The process of transitioning to a different cloud provider is known as cloud service migration. In any case, successful migration to a service provider‟s environment may require the use of

middleware, such as a cloud integration tool, to bridge any gaps between the vendor‟s and the customer‟s (or other vendor‟s) technologies.

Transitioning to the cloud or between cloud environments presents the usual IT issues, but the problems are compounded by having data stored and managed remotely, by external organizations and often in multiple locations. Among these issues are special considerations for privacy, interoperability, data and application portability, data integrity, business continuity, and security.

## REFERENCES

[1] Mr. Shrikant D. Bhopale,"Cloud Migration Benefits and Its Challenges Issue", IOSR Journal of Computer Engineering (IOSR-JCE) ISSN : 2278-0661, ISBN : 2278-8727, PP : 40-45
[2] Neetu Kishore and Seema Sharma,"Secured Data Migration from Enterprise to Cloud Storage – Analytical Survey",BIJIT -BVICAM's International Journal of Information Technology Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), New Delhi (INDIA)
[3] Rabi Prasad Padhy, Manas Ranjan Patra,Suresh Chandra Satapathy, "Cloud Computing: Security Issues and Research Challenges", IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS) ,Vol. 1, No. 2, December 2011
[4] Shyamli Dewan, Devendra Kumar, Sandeep Gonnade,"Secure Data Migration across Cloud System Using Third Party Auditor (TPA)", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 4, Issue 6, June 2015
[5] Noor Ibrahim Hussein, Mervat Hashem,"Security Migration Requirements: From Legacy System to Cloud and from Cloud to Cloud", 2nd International Symposium on Computer, Communication, Control and Automation (3CA 2013)
[6] Sultan Ullah,Zheng Xuefeng, "Cloud Computing Research Challenges", 2013
[7] Prashant Pant, Sanjeev Thakur,"Data Migration Across The Clouds", International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307, Volume-3, Issue-2, May 2013
[8] Wei Hao, I-Ling Yen and Bhavani Thuraisingham,"Dynamic Service and Data Migration in the Clouds", 2009 33rd Annual IEEE International Computer Software and Applications Conference
[9] Rashmi Rao, Pawan Prakash, "Improving security for data migration in cloud computing using randomized encryption technique", IOSR Journal of Computer Engineering (IOSR-JCE), ISSN: 2278-8727Volume 11, Issue 6
[10] Rajeshri Vaidya, and Prof. Sumedh Pundkar,"Large Data migration within Cloud Environments using Compression and Encryption Technique", International Journal of Innovative and Emerging Research in Engineering, Volume 2, Issue 2, 2015
[11] Mohammad Manzurul Islam, Sarwar Morshed and Parijat Goswami,"Cloud Computing: A Survey on its limitations and Potential Solutions", IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 4, No 2, July 2013