Auditing Cloud Computing Migration

Georgiana Mateescu*, Marius Vlădescu**, Valentin Şgârciu***

* University Polytechnics/Automatic and Computer Science, Bucharest, Romania
** University Polytechnics/Automatic and Computer Science, Bucharest, Romania
*** University Polytechnics/Automatic and Computer Science, Bucharest, Romania

*georgiana.mateescu@gmail.com, **vladescumariusnicolae@yahoo.com, ***vsgarciu@aii.pub.ro

Abstract – This paper presents a tool used to audit the cloud adoption within a company. The first section states the main advantages of the cloud computing environment together with its biggest challenges that can introduce significant risks within a company. In order to assess properly the cloud computing adoption, we developed a web based tool that assists the internal auditor in the pre-migration process. Section II describes in details the Migration Assessment Tool (MAT). Starting with its architecture, we presented the main objects manipulated by our application in order to implement a cloud adoption assessment. In order to compute the impact, we classified the questions in different domains and we defined the dependencies between the questions. At the end of section II, we presented the process implemented in MAT when assessing a cloud migration.

The paper ends with conclusions that state the main benefits of our approach and future improvements for MAT.

INTRODUCTION

Cloud Computing is a term that, although was first introduced in 1950s, it still inspires innovation, agility, development – future technology. The main reason for that is the fact that it can provide to a technology slave which is today business environment, an IT device independence from data availability point of view. Why is this important?

For sure, nowadays the information is more exposed than it used to be. Using the Internet and mainly thanks to distributed systems – the father concept of cloud computing [1], any data you need is at one search distance. The cloud computing architecture optimizes the IT infrastructure usage in order to maximize the organizations’ income and profit.

The main attributes that the cloud computing architectures provide are [2]:

- Shared resources – in cloud computing architectures multiple users utilize the same resources from network level, host level to application level.
- Massive scalability – because of its foundation principles, cloud computing has the ability to scale to thousands of systems.
- Elasticity – in cloud computing framework it is very easy to adapt both hardware and software resources to the user’s need.
- Pay as you go – cloud computing consumers pay only the resources they use just for the time they actually require them.
- Self provisioning of resources – additional systems (processing capability, software, storage) and network resources are added when and if they are needed.

These characteristics prove a lot of advantages including: lower-cost computers for users, improved performance, lower IT infrastructure costs for enterprises, fewer maintenance issues, lower software costs, instant software updates, increased computing power, unlimited storage capacity, increased data safety, improved compatibility between operating systems, improved document format compatibility, and universal access to documents [3]. There is also a list of challenges that can be even heavier than the benefits:

- This type of architecture requires constant Internet connectivity – moreover the disconnection can lead to a lot of inconsistencies that can be very hard to clean.
- The services quality is dependent on the connectivity power – it doesn’t matter how strong and reliable a service is, if for bad weather for example, the Internet is slower, the entire performance is compromised.
- The data ownership is shared between cloud provider and consumer which may lead to insufficient security in storing and using the data by the cloud vendor.
- Compatibility issues between the cloud providers which can lead to delays in processing data from systems hosted by different vendors, inefficiency in meeting the required SLAs.
- Regulatory compliance issues due to the geographical position of the cloud provider data center.
In order to take advantage of all cloud specific characteristics, the organization must first assess the systems suitable for Cloud migration. We propose an approach to quantify the impact of the cloud migration for existing applications within the organization architecture. Section II describes the Migration Assessment Tool that aims to offer to the decision taking actors within an organization a clear view on the existing options for cloud migration and the aspects that must be taken into consideration in such process.

**Migration Assessment Tool**

The Migration Assessment Tool (MAT) is a web based application that assists the enterprises in assessing the impact to migrate their IT architecture to cloud based environment. This system can be integrated also with an Identity Management solution in order to retrieve more information related to the applications that are being assessed for migration.

This section describes out approach for migration audit and addresses the following topics:

- **Architecture** – this sub-section describes the architecture of the MAT and the integration mechanism used to communicate with Identity Management system

- **Data model** – this sub-section describes the entities used in the MAT in order to manipulate the required data for an assessment process

  - **Impact Definition** – this sub-section describes the criteria used by MAT to assess the impact together with the domains of the each impact
  - **Assessment Approach** – this sub-section describes the process that is implemented within MAT when a migration assessment is conducted.

**A. Architecture**

MTA it is a web 3 tier based application with the following structure:

- **Presentation Layer** – this layer is implemented using HTML technology and represents the mean of interaction used by the end user to work with MAT. The application implements two roles: Administrator and Auditor. The full list of the access each role provides it is presented in the appendix.

- **Business Layer** – this layer implements the logic of MAT and manipulates the existing objects within MAT in order to perform specific operations such as: create/update/remove organization components, compute assessment and generate guidelines regarding available options for cloud migration, generating reports etc.

- **Data Layer** – this layer stores the data used by MAT. There are two main data sources: the MAT database and the Identity Management solution database. The Identity Management solution it is integrated using APIs in order to ensure accurate data related to the applications in scope for the assessment.

Figure 2 depicts the MAT architecture:

**B. Data Model**

In order to conduct a cloud migration assessment, MAT manipulates the several entities.

- **Enterprise** – this object is the representation of an enterprise within MAT. An enterprise it is composed by Clusters of Applications. Each enterprise can have defined an Identity Management solution that stores information related to the existing applications from the enterprise.

- **Identity Management Solution** – this object is the representation of an identity management solution within MAT. When defining a new Identity Management Solution (IdM), the user has to provide the credentials to the IdM solution together with the type of Database used
by IdM. The IdM is used only to retrieve information related to the application.

Cluster – this object is the representation of a set of applications that have common characteristics from cloud migration perspective. This kind of object can be managed only by MAT administrators. This means that for auditors, the cluster section will be a predefined list of values from where they will have to select the appropriate value when they create a new application. The cluster can be used as the scope of an assessment which means that the assessment will contain questions that have as target clusters. The assessment scoring will address the entire set of applications existing in the cluster assessed. Therefore when selecting a cluster as assessment scope, MAT will display a list of clusters that have defined at least one application for the specific enterprise.

Application – this object is the representation of the enterprise system that is assessed for migration. Each application belongs to a cluster. An application can be integrated with the Identity Manager solution defined for the enterprise – in this case the user will have to provide additional information related to the specific table from Identity Manager Database where the accounts of that application are stored. If the application is not integrated with IdM, the user will have to provide information related to the number of existing users from the application.

Cloud Provider – this object is the representation of a cloud provider within MAT. Each cloud provider has strengths and drawbacks and, depending on the application scoring it is evaluated as suitable or not suitable for an application migration. Each cloud provider addresses multiple deployment and delivery models that are taken into consideration during assessment.

Assessment – this object is the representation of an assessment process and can address a cluster or a single application. An assessment process consists of a set of questions with single choice selection based on which the assessment target scoring is computed by MAT. The questions are classified in Domains. Each assessment leads to an assessment scoring and an adoption plan. The assessment scoring in the process that, based on the answers provided by the auditor to the assessment questions, computes the score for each assessment target. There are different scores computed as part of the assessment, one for each delivery model. Based the lowest score – which means the lowest risk in case of cloud adoption, MAT suggests the most suitable cloud providers for the cloud adoption together with the strengths and drawbacks for each cloud provider.

Adoption – this object is the representation of an adoption process that is assessed by the auditor. The Adoption can address either an enterprise cluster, either a single application. This entity is made of tasks required in order to perform the cloud migration together with the risk scoring for each particular task. The scores are computed based on the Assessment performed for the adoption target. Therefore the Adoption object is a result of an Assessment Process. An adoption also includes best practices related to the suggested adoption plan.

The picture below depicts the relationship between MAT objects for assessment process.

Each assessment can contain one or more adoption, depending on the assessment targets. If the assessment target is a cluster, then the auditor can decide to perform an adoption for each of the application contained by the cluster or for the entire cluster. If the entire cluster is selected as target for the adoption, the referenced score for cloud providers evaluation is computed as the mean score of all applications within the cluster.

The picture below depicts the relationship between MAT objects for adoption process.
Each adoption has several tasks associated to it. The tasks are managed by MAT administrator and define high level processes that must be performed within a cloud adoption process. As a consequence, the list of adoption tasks is predefined and MAT associates them with the adoption plan based on the adoption scope and deployment model. Each task has a score associated to it that emphasize the risk level of the task. This risk level depends on the application specific information related to different domain considered during the assessment process.

C. Impact Definition

MAT performs the adoption assessment using an impact based algorithm. The impact represents the level of risk that the migration of a specific target presents from a certain perspective. The assessment process consists of a set of questions that the auditor must answer. One question can have the following scopes:

- Application – which means that the question will be part of the assessment process only if the assessment target is application
- Cluster – which means that the question will be part of the assessment process only if the assessment target is cluster
- Both – which means that the question will be part of all assessment processes

Each question has multiple possible answers, each of them having multiple scores associated to it. The scores associated to the answers are dependent on the deployment model. When conducting an assessment, the auditor must provide only one answer for each question. Based on all questions addressed by an assessment process, MAT computes the following scoring processes:

- public cloud score – this score represents the risk level for the assessment target if the target will be migrated to a public cloud. If this score is the minimum score obtained after the assessment process, MAT will search for the more reliable cloud providers for public cloud deployment models and it will retrieve them during adoption process. Depending on the specific cloud provider, MAT will provide a list of strengths and drawbacks for the cloud providers together with best practices for public cloud implementation.
- hybrid cloud – this score represents the risk level for the assessment target if the target will be migrated to a hybrid cloud. If this score is the minimum score obtained after the assessment process, MAT will search for the more reliable cloud providers for hybrid cloud deployment models and it will retrieve them during adoption process. When suggesting hybrid cloud adoption, MAT will provide also the list of sensitive items that must be considered for private clouds instead of public cloud. These items are retrieved using the highest scored questions for the public cloud score.

- private cloud – this score represents the risk level for the assessment target if the target will be migrated to a private cloud. Usually this score is minimum for the application with sensitive data because the best practices from cloud adoption recommend to keep the sensitive data within the company premises. As compared to internal dedicated, the data from the applications suitable to be migrated to cloud must be accessed also by external users.
- internal dedicated – this score represents the risk level for the assessment target if the target is not suitable to be migrated to cloud. Usually this score is minimum for the application with sensitive data that must be accessed only by internal users.

The picture below depicts the data model for the scoring mechanism implemented in MAT.
- $s_i$ is the score for the answer selected by the auditor during assessment process for the question $i$ which is not dependent on any other question.

If there is any dependency between the questions, the score is computed as following:

$$s_n = s_{n-1} + s_i \cdot q_{d_i}.$$  

Where:

- $s_n$ is the total score for the assessment that contains $n$ independent questions
- $s_{n-1}$ is the score for the first $n-1$ independent questions
- $s_i$ is the score for the current question that is dependent on question $q_{d_i}$
- $q_{d_i}$ is the score of the question which the current question is dependent on.

The impact represents the minimum score computed for an assessment:

$$i = \min(s_{pbc}, s_{hc}, s_{pc}, s_{id}).$$

Where:

- $i$ is the assessment impact
- $s_{pbc}$ is the score for the public cloud adoption
- $s_{hc}$ is the score for the hybrid cloud adoption
- $s_{pc}$ is the score for the private cloud adoption
- $s_{id}$ is the score for the internal dedicated approach

Depending on the impact, MAT suggests a deployment model and recommends the cloud provider more suitable for the deployment.

**D. Assessment Approach**

In order to conduct a cloud adoption assessment, MAT implements the following process:

1. Define Assessment Scope
2. Select target
3. Send questions and answers for target
4. Provide Answers
5. Compute Impact
6. Compute cloud selection criteria
7. Send suitable cloud providers and best practices
8. Prepare assessment report and adoption process

![Assessment process diagram](image)

All questions are classified into the following domains, depending on which aspect does it address:

- **Implementation Complexity** – this domain addresses challenges faced due to the complexity of migrating the business processes from on premise application to cloud application. These challenges are generated by the low degree of flexibility of the cloud application as compared to the large variety of business needs.
- **Risk and Compliance** – this domain addresses challenges in regulatory compliance. Usually the questions included in this domain are related to security aspects imposed by the existing regulatory in specific domains such as medical systems etc.
- **Infrastructure** – this domain addresses challenges when migrating the existing application from infrastructure perspective. In this domain are usually included questions related to specific features that the devices that host the application should have – such as Intrusion Detection System etc.
- **Performance** – this domain addresses challenges in ensuring the required performance for the migrated application. The questions included in this domain are related the Service Level Agreement that would be contracted with the cloud provider in case of cloud adoption.

After answering the provided questions, MTA performs the scoring and computes the impact of the cloud adoption. Based on the impact, MTA suggests the most suitable cloud providers and provides the criteria based on which the selection was performed.

The assessment results are presented together with the community best practice for the suggested deployment model.

**CONCLUSIONS**

“Cloud computing promises to have far-reaching effects on the systems and networks of federal agencies and other organizations. Emphasis on the cost and performance benefits of public cloud computing should be balanced with the fundamental security and privacy concerns federal agencies and organizations have with these computing environments.”[4]

In order to ensure that the risk raised by the cloud computing environment within an enterprise architecture is not greater than the adoption benefit itself, the enterprise must perform accurate audit process that quantifies the advantages and challenges a cloud migration inserts in the enterprise environment. In this paper we presented a practical tool used to assess the impact of cloud computing adoption. Starting from the best practices and cloud community experience, we listed out all relevant questions that an auditor should response when considering cloud adoption. Based on the domains addressed by each question, we scored each possible answer according to the risk inserted in the company. In this way, we were able to recommend the
best cloud provider for the specific environment and the most suitable deployment model.

The main advantages of our approach are:

- MAT provides a user friendly tool that assesses the impact of cloud migration according to the existing best practices from cloud community.
- By assessing the impact of the cloud adoption, the enterprises are able to quantify the risk implied by cloud environment.
- MAT provides a strong decision making support on long term strategy by providing a list of existing cloud providers suitable for specific environments together with theirs advantages and disadvantages.

In future, we plan to enhance the MAT to include "what-if analysis" that addresses dependencies between cloud and on premise applications. Also, another opportunity of improvement it is to include other module that audits the existing security mechanism within the enterprise application and include them in the algorithm used to compute the impact of cloud migration scoring.

REFERENCES


APPENDIX

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>OBJECTS MANIPULATED BY ROLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role Object Manipulated</td>
<td>Administrator</td>
</tr>
<tr>
<td>User</td>
<td>x</td>
</tr>
<tr>
<td>Enterprise</td>
<td>x</td>
</tr>
<tr>
<td>Role</td>
<td>x</td>
</tr>
<tr>
<td>Best Practice</td>
<td>x</td>
</tr>
<tr>
<td>Level</td>
<td>x</td>
</tr>
<tr>
<td>Cluster</td>
<td>x</td>
</tr>
<tr>
<td>Application</td>
<td>x</td>
</tr>
<tr>
<td>Question</td>
<td>x</td>
</tr>
<tr>
<td>Scope</td>
<td>x</td>
</tr>
<tr>
<td>Cloud Provider</td>
<td>x</td>
</tr>
<tr>
<td>Adoption</td>
<td>x</td>
</tr>
<tr>
<td>Identity Mgmt System</td>
<td>x</td>
</tr>
<tr>
<td>Assessment</td>
<td>x</td>
</tr>
<tr>
<td>Scorecard</td>
<td>x</td>
</tr>
<tr>
<td>Task</td>
<td>x</td>
</tr>
</tbody>
</table>