Cloud Interoperability via Quick Enterprise Applications Re-Builds

Nikolai Joukov and Vladislav Shorokhov

modelizeIT Inc., Stony Brook, N.Y., U. S. A.

Keywords: Cloud Interoperability, Enterprise Application.

Abstract: Cloud interoperability is a new problem that is becoming evident as more cloud providers offer competing clouds and more enterprises migrate their applications into the cloud environments. However, for large enterprises the problem lies in enterprise software and middleware non-interoperability and overall complexity more than in the cloud providers incurred issues. In this paper we analyze cloud interoperability issues for existing enterprise applications based on our substantial enterprise IT transformation experience. We believe that adoption of stricter enterprise application development and maintenance policies and their enforcement coupled with PaaS clouds' quick middleware provisioning and configuration capabilities will allow easy, predictable, fast, and inexpensive application rebuilds on-demand. This, in turn, will allow much cheaper application transformations, adoption of new technologies, and migrations between various clouds and non-cloud environments. We evaluate our policies based on three real-life applications from large corporations and show that strict application documentation and standardization results in at least an order of magnitude cost reduction of cloud application migrations.

1 INTRODUCTION

Cloud interoperability is a new problem that cloud users started facing now when a variety of cloud providers fiercely compete by lowering prices and offering new features. While it is much easier for small cloud users to flock from one provider to another, chasing new benefits, large IT clients have significant inertia due to their IT complexity and related complexities of IT changes. Vendor lock-down is a traditional fear of large companies and enterprises are trying to diversify their IT assets. However, the bigger problem is enterprise IT inflexibility in general. There is ongoing need to change IT according to constantly changing needs. Clouds, however, make infrastructure, management software, and management processes ever more tightly controlled by the cloud provider.

Fortunately, clouds, due to their uniform management and uniform policies, can make IT transformations and inter-environment migrations easier instead of making them harder. Therefore, it is critical for large companies to plan their cloud strategy in advance and plan for IT flexibility and not just fear a cloud provider lock-down.

IT interoperability is not "black and white": migration costs can be dramatically different for different types of IT transformations (Perng and Chang, 2012). Most enterprise cloud vendors are ready to even customize their clouds to accommodate large clients. Similarly, hard-to-migrate applications could be redesigned and re-implemented to work with a different cloud but the costs reflect the efforts and may become prohibitive even for very attractive savings opportunities post-migration. On the other hand, a simple test database migration from cloud vendor to vendor may consist of several mouse clicks to provision a new image with a database. Therefore, we may assume that any cloud is compatible with any other cloud but the cost of migration can be vastly different.

Any IT transformation is a service and lowering IT transformation costs by standardizing processes, methods, and by using automated tools is a focus of services science research (e.g., Bichler and Bhattacharya, 2011). In large enterprises IT infrastructure is typically large and complex with many unknown aspects of operation, inter-dependencies, and components inherited from decades ago or from other acquired enterprises.

Vast majority of enterprise applications today consist of operating system images with various interdependent software components. These applications rely on enterprise storage components, security mechanisms, and networking infrastructure. Not to mention that operating system images consume CPU, memory, and I/O resources.

Different data centers and portions of IT typically have different purposes, requirements, legacy-related restrictions, and composition. Therefore, it would be more appropriate to say that in a typical enterprise IT consists of several hybrid environments as show in Figure 1.
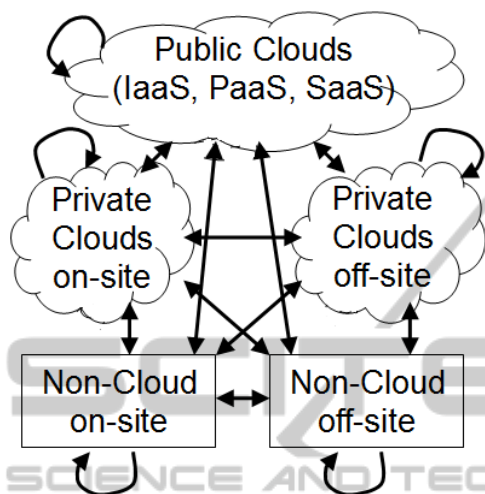

Figure 1: Enterprise IT environments are hybrid today and will likely stay diverse for many enterprises. Arrows show possible application migration scenarios that each requires cloud interoperability.

Hybrid enterprise environments including clouds exist for a number of reasons.

Non-cloud on-site: Some applications must stay with the client and even run on custom hardware for performance reasons (e.g., high frequency stock trading), some applications require tight control by the personnel (e.g., nuclear reactor control), some applications are too old and expensive to be modified or moved anywhere. All these applications are staying with the clients on-site.

Non-cloud off-site: Service outsourcing made it possible to move some client hardware together with applications to collocation data centers. Some performance-sensitive applications and large databases will likely continue to run on dedicated hardware without virtualization let alone clouds.

Private clouds on-site: Compliance and security concerns force enterprises to keep their applications on-site but still take advantage of clouds for easier management and resource provisioning.

Private clouds off-site: Some enterprise workloads run in remote data centers managed by trusted cloud vendors. However, clients like banks require dedicated network links from their offices to cloud data centers and completely isolated hardware.

Public IaaS clouds: Some non-sensitive enterprise applications and most development and test servers enjoy public clouds even today.

PaaS clouds: Enterprise applications rely on a number of popular components. Since they are standard they can be more efficiently provisioned and managed by a cloud provider.

SaaS clouds: It is likely that many standard applications in the future will be consumed as a service. (However, it is possible that a client will want to migrate back to a private solution again.)

Enterprise IT environments being hybrid now may stay hybrid forever. Technologies change, companies merge together and merge IT, and some applications become old, with no people deeply familiar with their operation. Moreover, there always going to be changes between cloud vendors, regardless of the price due to changing political preferences at CIO level and above, client dissatisfaction, new technologies, new internal cost factors, changing regulations, security requirements, and various other reasons.

Therefore, IT environments will be transformed and may require migrations along any arrow depicted in Figure 1. In this paper these arrows depict whole proprietary application migrations "as is" between environments and not interoperability between interfaces of composite applications (e.g., Hadar and Danielson, 2012; Hadar et al., 2012).

Migrations at the application-to-application level and not virtual image to virtual image level are a lot more flexible in terms of adoption for every new environment's features, restrictions, and benefits. Therefore, we will focus on this type of migrations.

In this paper we explore how today's typical processes to migrate applications from physical to virtual and to cloud environments could be extrapolated into future cloud to cloud migration processes (Section 2). We explore policies that, if implemented on the client side, may be even more useful and practical than on a cloud provider side (Section 3). We provide analysis of 3 real enterprise applications in the context of cloud to cloud migrations (Section 4). We conclude in Section 5.

2 RELATED WORK

IT transformation, including application cloudification, is a popular area of research in the recent years with many dozens of important papers published every year. In this section we describe a

typical cloud transformation process and refer to some recent underlying publications.

IT transformations follow a common pattern be it a transformation to a cloud, server virtualization, storage or network or security-related optimization, or some other IT infrastructure change (Pfitzmann and Joukov, 2011-1). Before proceeding to the next sections it is important to understand how IT transformations including, of course, migrations to clouds happen today (Ward et al., 2010). Figure 2 shows standard steps required.

Discover ▶ Plan ▶ Implement ▶ Test

Figure 2: General IT transformation process.

Before any transformation can begin it is necessary to collect information about the current situation with all IT components (discover), their utilization, and interdependencies (Bai et al., 2013). Thus, it is necessary to discover hardware, software, applications inventory (Joukov et al., 2008), business importance of the assets, and related risks they can tolerate (Joukov et al., 2009), aspects of storage (Joukov et al., 2010), and network operation (Ramasamy et al, 2011).

Based on the discovered information it is necessary to estimate the feasibility of the transformation, its costs, future benefits, timing, risks, and required resources. For large enterprise clients with dozens of billions in revenue per year and, sometimes, billion dollar a year IT budgets there is no such thing as impossible transformation. Both client applications can be adjusted and target cloud providers can adjust their clouds. However, the costs do matter. Easy to migrate applications get migrated today and others are predominantly still not in a cloud. A standard approach to estimating complexity is to consider client workloads, target cloud capabilities, and costs incurred by the specific transformation teams. Before assigning a real transformation cost value each application is assigned a complexity score (e.g., IBM GTS, 2011).

Transformation complexity depends on such factors as possibility of OS image importing to the cloud, which is today much cheaper than re-building new OS/application stacks (Assuncao et al., 2012). Clouds always restrict selection of OSs available to gain better benefit of uniform environment management. Today enterprise clouds allow import of at least Linux and Windows images from the clients (Amazon, 2010; Niijima, 2012). However, applications running on other platforms such as AIX, HP-UX, and Solaris are not as easily portable.

Nevertheless, even if an OS image import into the cloud is possible it may limit the benefits offered by the cloud. Enterprise clouds today provide middleware management such as patching and fast provisioning of middleware instances at least for sets of images with common middleware (Pfitzmann and Joukov, 2011). Therefore, in many cases, migration into a cloud happens as a complete application reinstallation and reconfiguration on new OS images and, sometimes, middleware installations offered by the cloud provider. Obviously this type of migration is much more expensive than image importing today especially if OS type, version, or middleware type or version change (IBM GTS, 2011).

Not every cloud or cloud image is suitable for every workload from the performance point of view. While CPU and memory measurements are easier, I/O-related benchmarking is especially challenging because multiple unknown machines compete for shared resources (Tak et al., 2013).

Enterprise applications are especially sensitive to survivability and disaster recovery guarantees. OS and middleware clustering between images and data centers is a norm. To support these requirements the clouds must support fast shared storage, special image monitoring software, ability to specify that images and data should not share same hardware and software or even separate them into zones like CloudStack does (Baset, 2012).

In addition to technical limitations, there are non-functional requirements such as security, various governmental regulations that limit possible data locations and operations on them.

Once a transformation process commences it is necessary to test that all aspects of the application operation are correct. This area includes classical systems testing problems (Ding et al., 2010).

Standard middleware platforms for enterprise applications such as Google's App Engine (appengine.google.com) require a complete rewrite of all legacy applications that enterprises have today and thus, related transformation is still prohibitively expensive. Popular IT standardization projects that let system administrators codify and recreate configurations such as Puppet (puppetlabs.com) and Chief (opscode.com) also require codification and re-configuration for existing applications. In addition, by default, they do not verify and enforce deep application-level configuration standardization.

3 PORTABILITY OF ENTERPRISE APPLICATIONS

Cloud providers do not typically create restrictions on the clients' ability to migrate applications–they simply cannot afford to provide fewer freedoms or restrict the clients more than their competitors. However, this freedom for the clients results in the situation that clients themselves abuse the freedoms and create unique and non-standard applications that are too expensive to migrate. The only real incompatibility issue that a cloud provider can trick a client into can be cost-driven or driven by unique new technologies. For example, a cloud provider can have special deals for software licenses that could not be moved to a different environment. Similarly, a cloud provider may offer a unique database engine at a discounted price compared to portable solutions with similar performance.

We believe that in all such cases cloud interoperability for enterprise applications, and IT transformability in general, can be achieved by adopting standardization policies by the clients.

3.1 Standardization Policies

We consider two key policies for enterprise applications aimed at their cloud interoperability:

- **Documentation:** Maintain application design, implementation, and change documentation;
- **Standardization:** Applications and their configurations should rely on common middleware and portable configurations.

These two policies can make rebuilding an application in a new environment a predictable and easy process and, thus, making it easy to rebuild any application in a new cloud easily.

Complexity of migrations from cloud to cloud (as well as complexity of any IT transformation) is hard to predict. Application owners, system administrators, and other administrators frequently make changes to address some immediate goal with the assumption that they know what they changed in their domain of responsibility. However, a massive transformation requires the knowledge of the current IT state as well as the reasons behind this state and past problems that were already solved. Keeping track of all changes and migration processes in the past and detailed description of the customizations required for each application can help predict future migration costs and complexities, and lower these costs. Essentially, each change can be quickly reproduced if it is documented and if it is not,

migration teams have to rediscover and reinvent existing solutions. It is essential that the application model and documentation reflect not just the status quo but also reasons behind decisions (e.g., for the type of clustering chosen). This can help quickly adjust the design for a new environment if the original design does not easily fit the new cloud. Documentation and application models should also contain all the typical information required for IT transformations like performance, reliability, security, compliance described in Section 2.

Cloud interoperability will obviously become challenging for the clients that decide to rely on proprietary cloud software such as unique databases. However, even common middleware-based solutions today are rendered non-portable by ad-hoc customizations via scripts and custom application code. Refusal from non-common customizations can make application rebuilding in a new cloud as easy as provisioning standard middleware and porting configurations and data using common tools based on documented processes.

Maintaining deployment scripts, up-to-date documentation, obeying standards has initial (C_{init}) and ongoing cost per year (C_{doc}). The benefit (B) is realized over a long period of time (t). If B_t is saved per year on average:

$$B = (B_t - C_{doc})t - C_{init} \qquad (1)$$

Enterprise clouds, with their ability to automatically provision and manage common middleware configurations, make application re-deployments an attractive and easy option. It opens up completely new opportunities in the areas of fast migration, disaster recovery, testing and adoption of new technologies. This, in turn, should drive the rate of transformations up, B_t and, thus, overall benefit from standardization and proper documentation up.

3.2 Policies Enforcement

No voluntary standards would be useful without enforcement methods. We foresee the following:

1) Periodic application rebuild tests based on documentation similar to disaster recovery tests conducted periodically today;
2) Voluntary authorization restrictions (e.g., give up root rights and application change capabilities to the cloud provider);
3) Automated discovery that can verify if documentation corresponds to reality and if only standard middleware and configurations are used (this type of verification is better performed by external audit teams).

4 EXAMPLES

A full-blown large-scale public evaluation of the proposed policies today is hardly possible because enterprise applications migration is a highly competitive business. All large-scale migration cost benchmarks are strictly confidential and represent a significant intellectual property.

We do our analysis based on three real-world enterprise applications that we have transformed, built, or optimized for large corporations. We have intimate knowledge of their operation and day-to-day management. Table 1 below lists basic properties of these applications: A1, A2, and A3.

Table 1: Analyzed enterprise applications.

	A1	A2	A3
Total servers	6	5	2
Test/QA servers	3	2	0
Cloud I/O performance is critical		+	
Massive interdependencies	+		
Hardcoded configuration			+
Geographic cloud zones required	+		
Network security zones required	+		
Relies on common middleware	+	+	+

A1 is used in a large enterprise to synchronize the operation of dozens of other applications, running on over a hundred servers, including their data exchange in multiple geographies and multiple network security zones. It is based on the classical three-tier architecture that consists of an IBM http server, an IBM WebSphere application server, and an IBM DB2 database. Because so many applications rely on correct operation of A1 it is geographically clustered at the application level and database data is replicated between these two geographies using database replication mechanisms. The application has a replica for tests and quality assurance purposes.

It is easy to notice that all middleware components used by this application are standard and can be provisioned with several mouse clicks in today's enterprise clouds. Moreover, tasks such as geographic zones creation and server assignment, network security zones provisioning, database replication configuration are partially automated even today and it is natural to expect complete automation of such tasks for enterprise clouds once more enterprise clients migrate their production environments. Today, most of these tasks are performed only during the application creation because of their complexity. However, these tasks in the clouds are much easier and could be repeated efficiently and inexpensively. For A1 specifically,

even the database content can be regenerated or migrated with a few commands. Based on our experience building enterprise applications (including A1) we estimate that if 1) A1 build process is completely documented; 2) all further customizations are documented and/or scripted; 3) common enterprise application provisioning-related operations above are automated and require just a few mouse clicks: the process of re-deploying A1 in a new environment from scratch will take under a day instead of about 10 days for a skilled engineer who has not deployed this specific application recently. Of course, this operation should be followed by application testing but that phase is required no matter how an application is migrated from cloud to cloud.

A2 is a standard SAP application with one server used as a user console, and two servers used for application part and its database (both replicated on test servers). A2 owners faced a performance problem after a change in the environment. It turned out to be related to database storage performance sensitivity. Unfortunately, the investigation took weeks instead of under a day because no performance data before the change was available for analysis. This again highlights the need for proper documentation of all aspects of application operation for smooth IT transformations.

A3 is an old application (that is typical for large enterprises that are in existence for decades). It consists of a Java EE application deployed on an application server. One of the goals of Java EE standard was to provide ease of application deployment and redeployment. Unfortunately, the standard is optional and majority of Java applications today are not designed according to the standard in a portable way. For example, application dependencies on other components, such as messaging queues and databases, are hardcoded in the application code. As a result, it may take migration engineers unprecedented efforts, sometimes measured in months, to migrate such applications today (Joukov et al., 2011). Proper use of Java EE standard or even detailed documentation that truly correspond to reality would make migration of such Java EE applications automatic.

As we can see, for all three applications proper documentation and standardization can cut the migration efforts by at least an order of magnitude given the PaaS clouds' ability of enterprise middleware provisioning and configuration.

5 CONCLUSIONS

Well-maintained documentation and standardization of IT assets is a dream of any CIO today. However, efforts required were always too high and benefits not well defined. In any case, IT operations relied on significant amount of human labour so any major IT change was always very costly. Clouds are about to change this situation dramatically by automating common operations for a large volume of clients. However, this is only possible if IT is standardized. Standardization *and proper documentation* for cloud applications open up a new opportunity: ability to re-build applications easily on demand. While it may not be a single mouse click it may still be cost efficient overall. Thus, rapid and inexpensive migrations between clouds, adoption of new technologies, or even new disaster recovery approaches will become possible.

Moreover, in the past it was not possible to even verify in a cost-efficient way if an application is well documented and relies on standards. We believe that mature discovery technologies and rapid test provisioning will become the basis of application standardization and documentation controls.

In this paper we considered three real-life enterprise applications and showed that documentation and standardization can make them easily re-deployable and inter-operable in modern enterprise clouds with effort reduction of at least an order of magnitude. Now, when enterprises are migrating their applications into the clouds, is the best time to adopt new policies as part of the migration process.

REFERENCES

Amazon, 2010. http://aws.amazon.com/about-aws/whats-new/2010/12/15/announcing-vm-import//

Assuncao, M. D., Netto, M. A. S., Peterson B., Renganarayana, L., Rofrano, J., Ward, C., Young, C., 2012. CloudAffinity: A Framework For Matching Servers to Cloudmates, *In 13th IEEE/IFIP Network Operations and Management Symposium (NOMS'12)*

Bai, K., Ge, N, Jamjoom, H., Zhang, X., Jan, E., and Renganarayana, L., 2013. What to Discover Before Migrating to the Cloud. In *IFIP/IEEE Integrated Network Management Symposium IM 2013.*

Baset, S. A., 2012. Open source cloud technologies. *In Proceedings of the Third ACM Symposium on Cloud Computing.*

Bichler, M., Bhattacharya, K., 2011. IT Service Management and IT Automation - Methods and Models for Efficient IT Operations, *In Business &*

Information Systems Engineering 3(1): 1-2 (2011) Article No. 28

Ding X., Huang H., Ruan Y., Shaikh A., Peterson B., Zhang X., 2010. Splitter: a proxy-based approach for post-migration testing of web applications. *In EuroSys 2010: 97-110*

Hadar, E., Danielson, D.J., 2012. Certified IT Services In-a-box for Cloud Computing Environments. *In CLOSER 2012*, 211-215

Hadar, E., Hadar, I., Ferguson, D.F., 2012. QDSL - Quality Domain Specific Language for Cloud Composite Applications. *In CLOSER 2012*, 228-233

IBM GTS, 2011. Get More out of Cloud with a Structured Workload Analysis,http://www-935.ibm.com/services/us/en/it-services/workload_transformation_analysis_for_cloud.html

Joukov, N., Devarakonda, M., Magoutis, K., and Vogl, N., 2008, Built-to-Order Service Engineering for Enterprise IT Discovery, *In IEEE International Conference on Services Computing (SCC'08).*

Joukov, N., Pfitzmann, B., Ramasamy, H. V., Vogl, N. G., Devarakonda, M. V., and Ager, T., 2009, ITBVM: IT Business Value Modeler, *In IEEE International Conference on Services Computing (SCC'09).*

Joukov, N., Pfitzmann, B., Ramasamy, H. V., and Devarakonda, M. V., 2010, Application-Storage Discovery, *In the 3rd Annual Haifa Experimental Systems Conference (Systor'10).*

Joukov, N., Tarasov, V., Ossher, J., Chicherin, S., Pistoia, M., Tateishi, T., 2011, Static Discovery and Remediation of Code-Embedded Resource Dependencies, In *IFIP/IEEE International Symposium on Integrated Network Management (IM'11).*

Niijima, T., 2012. Import a Linux OS to SmartCloud Enterprise., http://www.ibm.com/developerworks/cloud/library/cl-importlinuxOSimage/

Perng, C., Chang R., 2012. Methodology and Tool Design for Building Return on Investment Models for IT Transformations, *In 9th IEEE International Conference on e-Business Engineering (ICEBE).*

Pfitzmann, B. and Joukov, N., 2011-1. Assets for IT Transformation Services, *In 20th Annual Frontiers in Service Conference.*

Pfitzmann, B. and Joukov, N., 2011. Migration to Multi-Image Cloud Templates, *In IEEE International Conference on Services Computing (SCC 2011).*

Ramasamy, H., Tsao, C., Pfitzmann, B., Joukov, N., Murray, J.W., 2011. Towards Automated Identification of Security Zone Classification in Enterprise Networks, *In USENIX Workshop on Hot Topics in Management of Internet, Cloud, and Enterprise Networks and Services (Hot-ICE '11).*

Tak, B. C., Tang CQ, Huang H., Wang L., 2013. PseudoApp: Performance Prediction for Application Migration to Cloud, *In IM 2013.*

Ward, C., Aravamudan, N., Bhattacharya, K., Cheng, K., Filepp, R., Kearney, R.D., Peterson, B., Shwartz, L., Young, C.C., 2010. Workload Migration into Clouds Challenges, Experiences, Opportunities. *In IEEE CLOUD 2010:* 164-171