The Five-Step Program To Cloud!

Mahesh H. Dodani, IBM, U.S.A.

1   VIRTUALIZERS ANONYMOUS

"A twelve-step program is a set of guiding principles outlining a course of action for recovery from addiction, compulsion, or other behavioral problems. Originally proposed by Alcoholics Anonymous (AA) as a method of recovery from alcoholism … The method was then adapted and became the foundation of other twelve-step programs such as Narcotics Anonymous, Overeaters Anonymous, Co-Dependents Anonymous and Debtors Anonymous. As summarized by the American Psychological Association, the process involves the following:

- admitting that one cannot control one's addiction or compulsion;

- recognizing a greater power that can give strength;

- examining past errors with the help of a sponsor (experienced member);

- making amends for these errors;

- learning to live a new life with a new code of behavior;

- helping others that suffer from the same addictions or compulsions."
 Wikipedia, Twelve-Step Program

Before we start of our journey into the cloud, let us establish the ground rules and roadmap for our journey. As I described in my previous articles on cloud computing, the typical adoption path for cloud computing is a progression starting with centralize which focuses on reducing the number of data centers and infrastructure complexity, to consolidate which reduces redundancy and wastage of IT resources along with the space needed for the IT resources, to virtualize which separates the application workloads from the infrastructure and manages the virtualized environment, to automate which removes the manual processes of running the business by automating provisioning and through usage and metering services of the resources, and finally to optimize where the system can sense and respond to changing workload requirements and move these workloads to best-fit infrastructures.

Most enterprises have already spent time to move towards virtualization of their infrastructure – however, this move from a virtualization mindset to a cloud mindset is where most of the issues with adoption arise. It is not a straightforward jump from

virtualization to cloud, i.e. you can't just add automation and optimization to your virtualized environment and conjure up a cloud. The virtualization mindset is deep-rooted, and we need a program that helps virtualizers see the error of their ways, and provide them with a new cloud "code of behavior" that will allow them to free themselves from virtualization and realize the benefits of cloud.

Before we get to the program itself, let us look at the problems of the transformation from a virtualization mindset to a cloud mindset from the perspectives of the two key roles: the user and the administrator. We discuss these issues by laying out the progression that is made as these roles transition from a traditional IT infrastructure to a virtualized infrastructure, and finally to the cloud. This will allow us to see the major "chasm" between the virtualized and cloud mindsets. In general, we can characterize the traditional IT infrastructure services as providing direct access to the IT resources (compute, storage, network), the virtualized services as "swapping" a virtual machine for the IT resources, and cloud services as a complete separation of the service from the underlying IT resources that service it. As you can see from this generic characterization, the move from traditional to virtualization requires very little change in mindset – as we are just swapping a virtual machine for the actual IT resources being used. However, the move from virtualization to cloud requires a major shift in mindsets. In the following discussion, we will focus on development/test environments (this seems to be a common area for many organizations to get started on cloud computing) to describe the details of the progression into cloud computing.

For the developer/tester (our user of IT services), in the traditional environment they requested IT resources (CPU, RAM, and storage) on which they could build their development/test environments to do their work. In this traditional environment, they were responsible for managing and maintaining their environments (typically made up of an operating system along with a software stack representing the development platform or the middleware configuration on which the application they are developing will run on for testing purposes.) Therefore, the developer/tester is responsible for installing the needed software on the IT resources, keeping it up-to-date to conform to security, product and enterprise standards, and managing the saving and restoration of the environments. Depending on the size of the enterprise and the developer/tester organization, the IT resources are typically managed directly by the developer/tester group or in the enterprises' data center. When moving to a virtualized environment, the major change is the swapping of the requested IT resources with a virtual machine. The major impact this move has is that the developer/tester is no longer required to manage the IT resources and operating system – as it is now the responsibility of the data center to manage these and keep them up to date with security and enterprise standards. However, note that the developer/tester is still entirely responsible for building, maintaining and managing their development/test environments. The move to cloud changes the entire dynamic. Developers/testers must be "re-programmed" to think entirely in terms of their development/test environments and not in terms of the underlying IT resources that will service their needs. They must define common development/test environment configurations, the services that they need (e.g. save/restore, clone) to perform on these
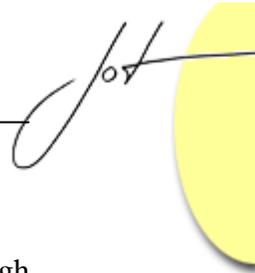
environments, and the qualities of service (e.g. for disaster recovery, data retention, high availability) that they require for these environments.

The administrator in a traditional IT environment focuses exclusively on monitoring and managing the IT resources (compute, storage, and network.) They allocate these IT resources based on application needs and the associated qualities of services requirements, monitor the allocated IT resources (and the overall data center which houses these resources), and manage the IT resources to deliver on the qualities of services that they have agreed on (including security, performance, high availability, disaster recovery, etc.) Furthermore, they manage the IT resource capacity that is available to support the applications, as well as plan to acquire capacity based on projected usage and requirements. In the virtualized environment, the administrator adds on the monitoring and managing of the virtual machines that they have allocated for use on top of the physical IT resources in their data centers. They now have the added responsibility of creating, configuring, allocating, monitoring and managing virtual machines. They also have to monitor, manage and plan for the capacity required by the virtual machines. In essence we have exacerbated the problems that the administrators faced with monitoring and managing physical IT resources by adding many more virtual machines to deal with. Another way to look at this is that by virtualizing we help get rid of server sprawl, but are replacing it with the larger (in terms of quantity) virtual machine sprawl. Even for the administrator there is a large jump when moving from virtualization to cloud computing. They have to change their mindset from focusing only on IT resources (and their virtual counterparts), to thinking about cloud workloads (e.g. development/test environments) along with how to monitor and manage these workloads. This move to the cloud mindset affects all the practices that they are used to as well. For example, to monitor and manage capacity needed to support the cloud workloads, administrators need to transform their thinking from acquiring further IT resources to acquiring cloud services (infrastructure-as-a-service) and move workloads to be handled by these services. Furthermore, the decisions on which cloud services to acquire and which workloads to move will be dependent on several criteria – workload qualities of services requirements, cost of cloud services, cost of moving and managing workloads, etc.

It is important to note that the very reasons that are touted as drivers for cloud adoption can be stymied if this mindset transformation does not occur. Lets take one example – increasing the utilization of IT resources. This is one of the key drivers for cloud adoption. However, in the virtualization mindset it is very difficult to realize increases in utilization, because swapping a virtual machine for physical IT resources does not result in any increases in utilization. Transforming into the cloud mindset breaks this direct association between services requested and the physical IT infrastructure which can lead to increased utilization. This transformation is needed in both the user to ensure that their usage patterns of IT resources changes only when they are actually doing work, as well as the administrator to ensure that they are allocating appropriate IT resources to support the users needs.

2   THE FIVE-STEP PROGRAM

Having understood the problems associated with cloud adoption, we can set up a program that can tackle the issues associated with the virtualization mindset, manage the transition to the cloud mindset (along with the associated behavioral changes), and help with keeping the organization entrenched in the cloud. Figure 1 summarizes the Five-Step program for cloud adoption. Step 1 is establishing the roadmap for transforming the IT infrastructure – moving from the traditional data center approach to a centralized and consolidated infrastructure, through virtualization and automation, and finally realizing an optimized IT infrastructure capable of delivering on the promise of cloud. Step 2 establishes the architecture that will provide the capabilities needed to deliver cloud services effectively and efficiently, considering the service consumer, provider and creator. Step 3 focuses on analyzing the workloads that are feasible to move to the cloud. Step 4 focuses on determining the right mix of delivery models to deploy and use the cloud services. Finally, Step 5 lays out the implementation approach for deploying the cloud services. Note that each step focuses on the transformation of the entire enterprise into the cloud mindset. Let us discuss each step in detail.
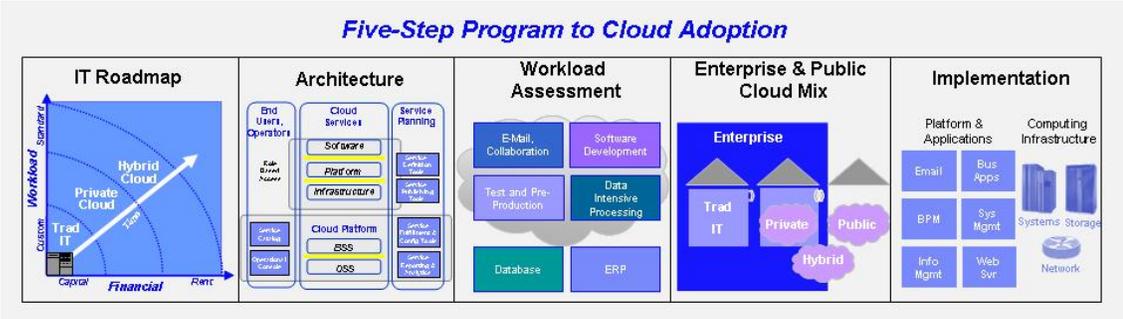

Figure 1: Five-Step Program to Cloud Adoption

Step 1 sets up the roadmap to transform IT, moving it through centralization and consoloditation, into virtualization and automation, and finally into an optimized cloud environment. Along this transformation, the enterprise must focus on implementing the required capabilities. In centralization and consolidation the transformation focuses on reducing infrastructure complexity and the staffing requirements on managing the infrastructure, improving business resiliency by managing fewer things better, and reducing the total cost of ownership (TCO) by improving operational costs. This is done by first establishing an enterprise data center strategy that aligns with the business needs, continuity requirements and geopolitical considerations, and then implementing the strategy including site relocation, consolidation, and new construction. The focus of consolidation is to move underutilized IT resources into larger, denser, scalable clusters, create pools of resources which can be used to service workloads, and managing and controlling these pooled resources. The next phase of the roadmap focuses on virtualization and automation. Virtualization reduces hardware costs and simplifyies deployments. It accomplishes this by defining virtual resources to separate physical IT
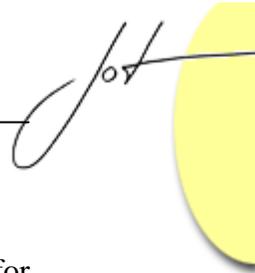
resources from its use to deliver services, establishing a single management system for virtual resources, integrating security and workload management, and controls virtual resources based on application requirements and Service Level Agreements (SLAs.) Automation focuses on dramatically reducing deployment cycles and enabling new processes/services through flexible delivery. It does this by reserving resources for applications through standardized images, automatically provisioning and de-provisioning resources based on reservations, managing workloads with integrated security and information virtualization. An optimized cloud environment will be able to sense workload requirements and move it to best-fit infrastructures while at the same time delivering qualities of service most cost effectively. It does so by optimizing workloads to maximize performance and efficiency, prioritizing workloads to attain SLAs, moving workloads to appropriate virtualized infrastructures to reduce costs, defining policies for workload management, and orchestrating workloads based on these policies.

Step 2 focuses on the cloud architecture, and provides a guide to understanding the capabilities required for an enterprise to derive value from implementing cloud. The architecture follows the separation of concerns priniciple, and subdivides the capabilities of cloud into the three main "parties" – the consumer of the cloud service, the provider of the cloud service, and the creator of the crowd service.

The cloud service creator component has capabilities targeting all aspects of the lifecycle of the "image" that is used to bundle the service that is accessible by the consumer. This "image" can include the IT resources (e.g. server, storage, network), the operating system, the middleware, and the applications. The capabilities support the need to design and build the image, store it in the library of images that can be accessed by the users, the deployment of the images, and the management of the images through the entire operational lifecycle.

The cloud service consumer serves both the end users and the operators that manage the infrastructure. The capabilities in this component ensure that the images that can be accessed are defined in a catalog, and have appropriate role-based user interfaces to access and manipulate the images.

The key capabilities of the reference architecture are defined in the service provider component. The lowest layer of the architecture defines the capabilities of the virtualized infrastructure. These capabilities facilitates virtualization of all IT resources: server, storage and network. These virtualization capabilities can handle all types of IT resources, e.g. both mainframe and distributed servers. The next layer provides an optimized middleware with capabilities for image deployment, integrated security, workload management, and high-availability. The optimized middleware is used as the way to deliver services and information built according to well defined SOA and Information architectures. The central piece of the component is service management which provides the capabilities to manage a cloud service. These services include capabilities to handle user requests: managing the self service requests made by users, the lifecyle of images, and the provisioning of images based on the request. The capabilities

also handle many of the qualities of service associated with delivering images, including availability, backup and restore, security and compliance, and performance management. To facilitate delivery through flexible models, the capabilities also support usage accounting and license management.

The reference architecture provides a comprehensive set of capabilities to ensure that cloud services can be built, deployed, accessed, delivered and managed. Each of these capabilities are supported by the appropriate standards, technologies and tools – all integrated to work together to deliver cloud computing.

Step 3 analyzes workloads that can run effectively in a cloud environment. It is important to understand the characteristics of the workloads in the context of a cloud delivery to determine its suitablity to be delivered as a cloud service. We want to avoid workloads where risk and migration cost may be too high, including those that are database intensive, have complex transaction processing, packaged application (e.g. ERP) workloads, and highly regulated workloads. We want to focus on workloads which can be standardized and take advantage of running as a cloud service, including web infrastructure applications, collaboration infrastructure, development and test workloads, and high performance computing workloads. Finally, we should consider new workloads that have been made possible by cloud including high-volume low cost analytics, collaborative business networks, and industry focused "smart" applications

Through our current experiences, we find the following workloads as feasible for current public cloud capabilities:

- Development, test and pre-production systems.

- Mature packaged offerings, like e-mail and collaboration.

- Batch processing jobs with limited security requirements.

- Isolated workloads where latency between components in the application is not an issue.

- Infrastructure as a service, including compute resources and storage as a service.

- Managed services for the IT infrastructure, including backup and restore, virus scans, etc.

Of course, our experiences have also shown that we should (with the current maturity of cloud) avoid the following workloads:

- Workloads which depend on sensitive data normally restricted to the enterprise, e.g. employee information, or health care records.

- Workloads composed of multiple, co-dependent services, such as high throughput online transaction processing

- Workloads requiring a high level of auditability, accountability, e.g. those subject to Sarbanes-Oxley.

- Workloads based on 3rd party software which does not have a virtualization or cloud aware licensing strategy.
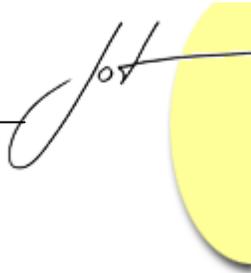
- Workloads requiring detailed chargeback or utilization measurement as required for capacity planning or departmental level billing.
- Workloads requiring extensive customization.

Step 4 focuses on determining how to effectively deliver the cloud services as shown in Figure 2. The cloud services can be delivered on premise or off premise using multiple mechanisms – either through traditional IT services or through managed operations. The financial models can include considerations for delivering the services privately, publically or through a hybrid model. Of course, different models can be chosen based on workload characteristics and the quality of service requirements. For example, development and test workloads that do not contain sensitive data, have low security requirements, and have low qualities of service requirements can best be serviced by public cloud services.
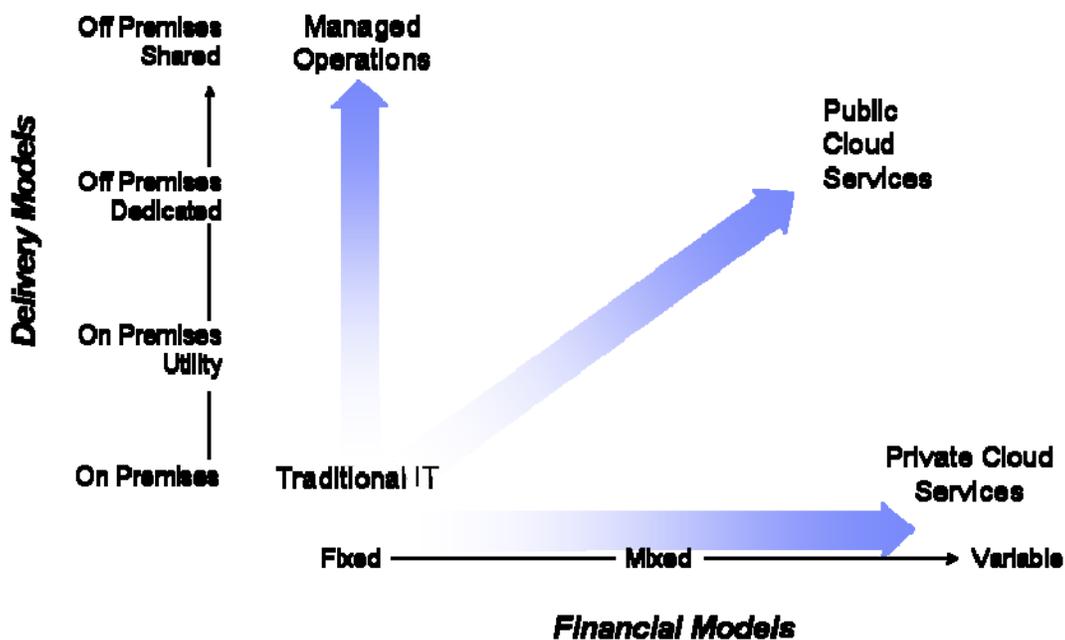

Figure 2: Cloud Services Delivery Models

Step 5 focuses on the implementation and deployment of the cloud service to derive business value for the enterprise. Each cloud implementation must focus on providing the capabilities for the service consumer, the service provider, and the service creator as were articulated in the cloud architecture described in Step 2. The following are key considerations in any implementation:

- The implementation should provide an easy to access, easy to use service catalog that is used by users to request services, and by administrators to publish available services.
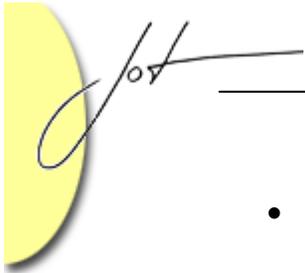
- The implementation should hide the underlying complex infrastructure from the user and shift the focus to services provided.

- The implementation should enable the ability to provide standardized and lower cost services.

- The implementation should facilitate a granular level of services metering and billing.

- The implementation should ease complexity through workload standardization.

The implementation should consider use cases for both users and administrators, ensuring capabilities for requesting/using/maintaining/releasing cloud services and for monitoring/managing/maintaining the cloud services and the underlying virtualized infrastructure.

3 GETTING A JUMPSTART

As the saying goes – every long journey starts with a single step. The same is true for the complex and difficult cloud journey that enterprises will embark on. Figure 3 shows the main ingredients and considerations for taking your first and very important step into cloud, where you plan and prepare for your journey into cloud, test and deploy with a well understood workload, and then extend and evolve into your cloud implementation.
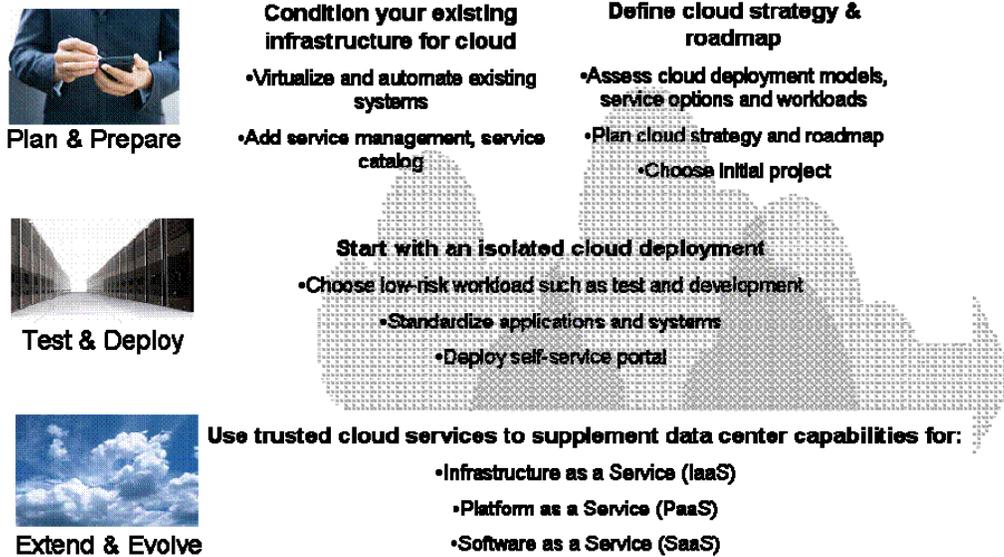

Figure 3: Getting Started on Cloud

Of course, a key first step is to get your virtualized infrastructure in place – at least the infrastructure that you will use to support your cloud services. Besides this tactical focus, it is also important to look at the overall cloud strategy and roadmap that will address your enterprises' needs. This will include key considerations as you walk through the

Five-Step program we outlined in the previous section. Once you have planned and prepared for the cloud, it is important that you identify and choose a workload that is reasonable for your first cloud deployment. The characteristics that are important as considerations for your first workload include that it is suitable for delivery as a cloud service requiring little or no modification (see Step 3 discussion in the previous section), it has minimal qualities of service requirements, and you can define the business value that will be realized in having it available as a cloud service. Once you have been successful at delivering this first workload and realized the business value, you can evolve and extend this experience into a full cloud deployment. In these phases, you will be involved as well in transforming your organization, governance, and business models to ensure that you can derive maximal benefit from your cloud implementation. You will also consider the right mix of delivery and financial models for deploying these cloud services as we discussed in Step 4 in the previous section.

In this article, we have set up our journey into cloud computing by establishing the pre-conditions, considerations and roadmap for a successful journey. Over the next few columns, we will continue our journey by taking a deep dive into the architecture and technical capabilies provided by cloud computing.

About the author

Mahesh Dodani is a software architect at IBM focusing on Cloud Computing. His primary interests are in enabling communities of practitioners to design and build solutions that address complex business needs and deliver value. He can be reached at dodani@us.ibm.com.