

Managing the Cloud as an Incremental Step Forward

How Service-now.com brings cloud services into your IT
infrastructure in a natural, manageable way

white paper

Table of Contents

Accepting the Cloud, Rejecting the Revolution	1
The Three Stages of Cloud Adoption	1
Stage 1: Virtual Infrastructure - My Hardware is Now Software	1
Stage 2: Dynamic Applications - Writing Applications for an Elastic World	3
Stage 3: The Responsive Data Center - Putting Workflows in Charge	4
How Service-now.com Tames the Cloud	4
Capabilities for Management of Cloud Resources	5
Capabilities for Support of Dynamic Applications	6
Capabilities for Support of Responsive Data Centers	6
Summing Up	7

Managing the Cloud as an Incremental Step Forward

Accepting the Cloud, Rejecting the Revolution

The cloud is revolutionary but with proper planning it won't create a mess.

Much analysis of what cloud computing means for business gets pretty excited pretty fast. Terms like “revolution,” “disruption,” and “paradigm shift” are thrown around wildly, as the analyst explains that everything about IT as you know it must change.

At Service-now.com, we think it is time to pause and reflect on two issues:

- What will the cloud mean for business in the short-, medium-, and long-term?
- How can IT harness the power of the cloud and incorporate it in an orderly manner?

Make no mistake; we see the cloud as a phenomenon that will change IT forever. After all, our product is built and based in the cloud. But the cloud can be revolutionary without being problematic only if cloud resources are incorporated into IT infrastructure in a way that allows them to be managed and controlled. Otherwise, the revolution will be a huge headache. Already, companies that have lurched into the cloud have created terms like “virtual sprawl” to describe the mess they have created.

To gain the obvious benefits offered by the cloud, organizations should have a clearly defined strategy for using and managing the cloud. This white paper analyzes two topics crucial to successful cloud exploitation:

- The lessons Service-now.com has learned from its client base about the stages of cloud adoption.
- How the capabilities of Service-now.com can automate, manage, and monitor cloud resources to make the introduction of cloud assets productive and smooth.

The Three Stages of Cloud Adoption

Service-now.com's unique view of data center Cloud usage has helped construct a model that describes successful patterns of adoption.

Service-now.com has the good fortune of having hundreds of the world's largest customers. Unlike other products in the IT service management space, Service-now.com's architecture allows us to get an intimate view of how our customers are using their assets (see “The Biggest IT Laboratory in the World”). Based on what we have seen at our clients—both big and small, early adopters and pragmatists—the following model of cloud adoption has emerged.

Stage 1: Virtual Infrastructure — My Hardware Is Now Software

The first stage of Cloud adoption replaces physical assets with virtual ones.

The first stage of cloud adoption replaces physical infrastructure with virtual infrastructure. A layer of software replaces what formerly was hardware, creating a virtual instance. Because the instance is software, it is changeable and controllable in ways that hardware is not.

It is important to understand that the concepts and technology of virtualization have been around for decades. IBM has been selling virtual machine hypervisors since the 1970s. VMware was founded in 1998 and led the early wave of deployment of virtualization technology in data centers and on the desktop. Today all major IT vendors offer virtualization or cloud products. The most common use of virtualization is to implement a program of server consolidation in which the number of physical machines is reduced using virtualization technology. Increasingly, virtualization is being expanded to

create private clouds, where virtualized assets are offered to internal users in an on-demand fashion.

Cloud computing is an extension of virtualization to the public network. Understandably, most of the action in the cloud has been focused on the low-hanging fruit, the use of cloud resources to provision basic infrastructure for computing power and various forms of reusable application plumbing like data stores and queues. The use of cloud assets in this way is most often called Infrastructure as a Service or IaaS.

Amazon Web Services has been the leader in the commercialization of such services, but many competitors such as the Rack-space Cloud and Navisite are also offering cloud computing services.

The benefits of the cloud come from the following attributes of cloud resources:

- They can be rapidly provisioned in minutes, instead of weeks or months
- They can be automatically provisioned through APIs
- They are elastic, meaning that you can expand and contract your use of resources as needed
- The cost is variable, meaning you only pay for what you use

The properties of the cloud have led to several common use cases:

- On demand and self-service creation of test and development environments
- On demand and self-service creation of environments for training and demonstrations
- Accelerated processing of high performance computing workloads by rapidly provisioned gangs of virtual machines
- Application load testing using rapidly provisioned gangs of virtual machines

The ease of use has been a tremendous benefit but has also led to a variety of problems:

- Cloud resources are sometimes used as a way to avoid working with internal IT departments, a strategy which seemingly provides acceleration until the need arises for professional management of the IT assets created
- Cloud resources can grow wildly out of control and lay unused
- Cloud resources, if not properly managed, can become a security compliance issue
- Cloud resources are frequently not properly backed up

The Biggest IT Laboratory in the World

One benefit of the wide adoption of Service-now.com is the creation of the world's largest source of data on how companies use IT assets. Because our product is based on the Software as a Service (SaaS) model, that is, hosted in the cloud, we are able to recognize macro trends in how our customers are using both our product and underlying IT assets. Service-now.com uses a small piece of code, called the MID Server, that runs inside the firewall of our customers. The MID Server acts as the eyes, ears, and hands of our SaaS application, allowing our clients both to monitor the assets of a data center and also to take action.

Our intimate view of these data centers comes from the usage patterns of our SaaS application as well as customer inquiries about how to push the envelope. As a company, we feel we are closer than anyone to the pulse of what is happening in the data center world-wide. That knowledge is the source of our insight into the cloud.

Some critics of the cloud see its downsides as fatal. At Service-now.com, we do not. The cloud is simply another class of IT assets that must be properly managed. The cloud is now enjoying a burst of attention just as end-user enthusiasm for spreadsheets, desktop databases, and the Internet represented a challenge to IT when first introduced.

Stage 2: Dynamic Applications — Writing Applications for an Elastic World

The second stage of cloud deployment changes the way applications are written.

Once a company becomes comfortable with IT assets, application architectures start to change to incorporate the flexibility of cloud assets. In stage 1 of cloud computing adoption, most of the assets are used as virtual replacements for physical assets. Virtual assets allow increased automation when creating assets and modifying configuration, but for the most part, the applications using virtual assets run as they always did.

In stage 2, applications start to monitor the flow of work coming through them. When the workload increases, the application uses cloud APIs to clone parts of itself and then spreads the work over the expanded infrastructure. Figure 1 shows how this architecture works.

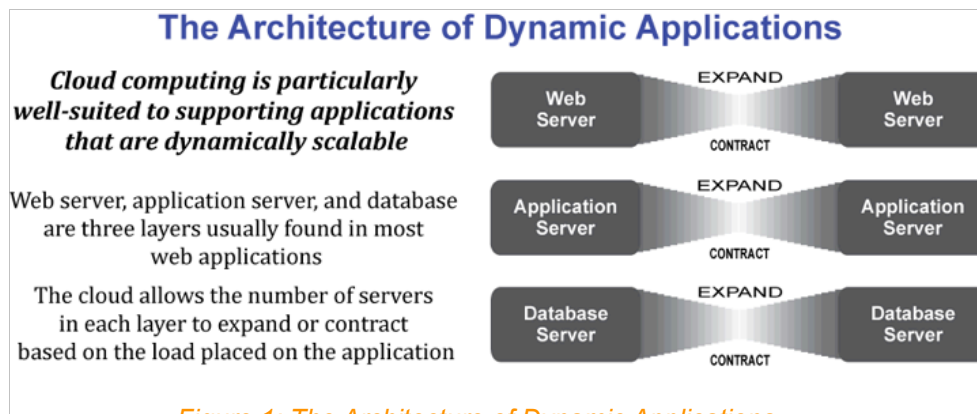


Figure 1: The Architecture of Dynamic Applications

This architecture has been wildly successful for applications like Animoto, a video creation application that became popular on Facebook, and for the online games created by Zynga. Web sites that have to handle large bursts of traffic also use this technique to dynamically expand and contract computing capacity as traffic grows.

A company called RightScale was one of the first enablers of these applications. RightScale’s technology, which is comparable to Service-now.com’s offering for run book automation, allows scripts to be created that can automatically create virtual machines, install the needed software, and bring them into production. Through a combination of monitoring inside the application and scripting outside the application, computing capacity can expand and contract dynamically.

The economic benefits of this approach are quite profound. It is costly to build large data centers in advance of the arrival of large amounts of traffic. But using a dynamic application architecture, you don’t have to. When your marketing succeeds and the traffic arrives, you can provision a large computing capacity when it is needed, pay for it only for a short while, and then contract the capacity as traffic slows. Animoto’s capacity grew from hundreds of servers to many thousands and back down again as its marketing succeeded.

While these examples of dynamic applications are compelling, the IT industry is just beginning to understand their implications.

Stage 3: The Responsive Data Center — Putting Workflows in Charge

The third stage of cloud adoption transforms the structure and workings of the data center, dramatically increasing automation and responsiveness.

Stage 3 of cloud adoption is only in place in companies with the most challenging computing environments. Such environments can be seen in financial services, oil & gas, and online advertising, which all feature rapidly expanding and contracting workloads that must be processed instantaneously.

In stage 3, the ability to expand and contract is applied not just to the application, but also to the data center and all its resources (see Figure 2). Making this work requires an executive function that has a model of all of the resources for computing, storage, databases, applications, networking, and so on for the entire data center.

The responsive data center is a complex and challenging entity to create and has been made possible by the arrival of so much virtualized infrastructure. In most responsive data centers, the virtual assets are internal, but as time passes, it is likely that assets in the public cloud will increasingly play a role.

The executive function monitors workloads coming in from various sources and then expands and contracts the data center and all applications involved in processing the workloads.

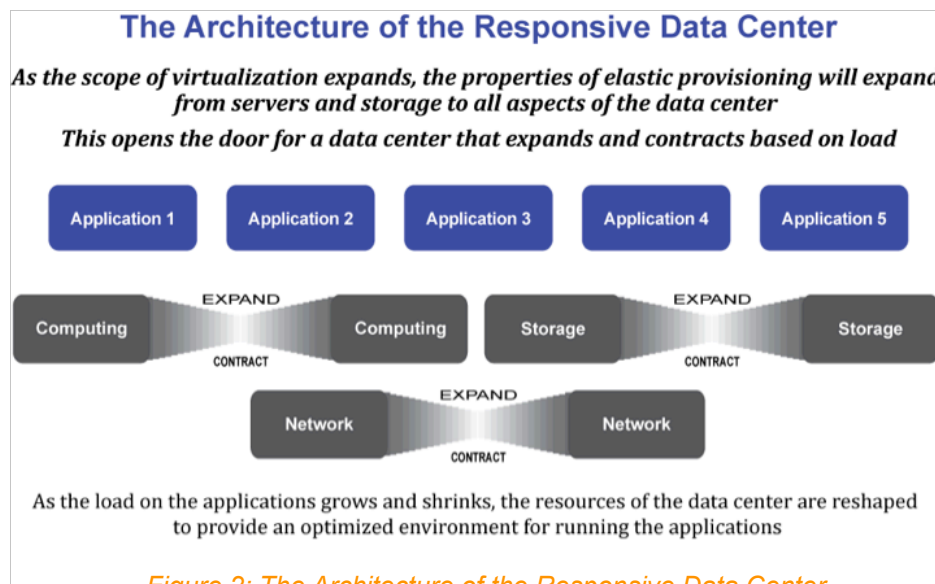


Figure 2: The Architecture of the Responsive Data Center

The responsive data center is a complex and challenging entity to create and has been made possible by the arrival of so much virtualized infrastructure. In most responsive data centers, the virtual assets are internal, but as time passes, it is likely that assets in the public cloud will increasingly play a role.

How Service-now.com Tames the Cloud

Service-now.com has a full suite of IT management capabilities for cloud resources.

Service-now.com’s view is that IT infrastructure, wherever it resides, should be managed in a unified and consistent manner. Service-now.com’s product development philosophy is based on observations of its clients’ use of IT assets and their

rapid response to requests for support of emerging needs. At our Knowledge user conference, we presented a list of credits showing how our customers inspired our application features, which are listed in Figure 3.



Figure 3: Service-now.com Customer-inspired Innovation

For quite some time now, we have been addressing our client’s needs to manage cloud resources. The next sections summarize how our capabilities support each stage of cloud adoption just described.

Capabilities for Management of Cloud Resources

From the Service-now.com point of view, a cloud resource is just another configuration item or service that is recorded in the configuration management database (CMDB) and that is managed and monitored through our capabilities.

Right now, Service-now.com can monitor and deploy virtual machines created inside the data center and virtual machines created on Amazon’s EC2 platform.

Here are some specific capabilities:

- Cloud resources are tracked in the CMDB
- Cloud resources are monitored through MID servers using cloud monitoring technology

- Cloud resources can be presented to users as part of the service catalog
- The creation and retirement of cloud assets can be automated using Service-now.com's run book automation
- Cloud resource providers and agreements are managed via service portfolio, and SLA, project, and capacity management

In a sense, cloud assets are like any other assets. They are computing resources used to do a job and that must be monitored and managed through a lifecycle. Cloud resources are different in that they are controlled and monitored through the APIs made available by the cloud provider. There is really no way to get under the hood.

The good news is that the web services APIs provided by Amazon and other cloud vendors provide a secure way to monitor and control cloud assets throughout their lifecycle. By controlling cloud assets through a set of standard scripts, it is possible to make sure that policies for security, backup, and management of sensitive data are followed.

On top of the primitive web service APIs, Service-now.com allows the creation of push-button functionality to set up a new computing environment. For example, attendees at Service-now.com's Knowledge10 user conference used dynamically created application sandbox environments during training classes. One click started a sequence of commands that created the test environment needed for a class. When the class was completed, the environment was deprovisioned with a click. In this way, cloud resources can be a welcome addition to a company's service portfolio.

Capabilities for Support of Dynamic Applications

Dynamic applications described above are supported through Service-now.com's run book automation capabilities, which allow scripts to be created to automatically provision and retire cloud resources.

The kind of support needed for dynamic applications depends on the scalability architecture. Figure 2 shows the most common three-layer architecture consisting of a web server, application server, and database server.

The support that dynamic applications need is as follows:

- The ability to recognize when the application is under load and needs to expand capacity
- The ability to create a new web server, application server, or database server to help share the load
- The ability to bring the new servers into production
- The ability to recognize when the load has dropped and the new servers are not needed
- The ability to deprovision servers that have been taken out of production

Service-now.com's customers have long relied on the company's ability to recognize that disks are close to filling up or that CPUs are under a heavy load. The same sort of monitoring capabilities and run book automation scripts can be employed in support of dynamic applications to perform all the functions listed above or any others that are needed.

Capabilities for Support of Responsive Data Centers

The most important capability for a responsive data center is an executive function that is monitoring all of the workloads coming in and deciding how to reconfigure that data center. Once a new form for the data center has been determined, scripts are run to perform the reconfiguration. Most of the time the reconfiguration is automatic, although it is possible that some types of reconfiguration may have to be performed manually if programmatic control of the resources is not possible. The quality of manual work can be assured using change management technology offered by Service-now.com to ensure risks are understood and that activities are correctly performed.

Service-now.com's CMDB is a huge asset to the executive function of a responsive data center. Combining the CMDB with run book automation provides a powerful tool for implementing efficient data center operations by creating higher-level actions that can be taken to perform the reconfiguration. Once changes are applied using run book automation or manually, Service-now.com's Discovery application can automatically verify that changes were executed as planned.

Summing Up

At Service-now.com, we believe the cloud is revolutionary, but it doesn't have to be a problem. With the monitoring, management, integration and automation capabilities of our product, the cloud can be deployed opportunistically without creating a mess. While other SaaS vendors look to take the power of the cloud away from IT by hiding its power inside applications, Service-now.com is helping to give it back by making it possible to leverage each type of resource for maximum benefit without creating a mess. We will help you bring the advantages of the cloud to your data center.