

Realtime  
publishers

*The Definitive Guide™ To*

# Cloud Computing

*sponsored by*



*Dan Sullivan*

---

Chapter 4: How Cloud Computing Will Help Your Business.....	62
How Cloud Computing Can Help Your Business .....	63
Business and Technology Alignments: The Ideal vs. Reality .....	63
Identify Business Priorities .....	65
Identify Operational Inefficiencies.....	66
Identify Barriers to Innovation.....	67
Assessing Current Capabilities .....	68
Infrastructure Capabilities .....	69
Platform Capabilities.....	69
OS Consolidation .....	70
Application Stacks .....	70
Application Capabilities .....	73
Governance Capabilities .....	74
Management and Reporting Capabilities .....	75
Introducing a New Model for Consumption and Delivery .....	76
Introducing Public Cloud Consumption Model.....	76
Introducing Private Cloud Consumption Model.....	76
Deploying Existing Infrastructure in a Private Cloud .....	77
Enabling Application Services in the a Cloud .....	78
Managing a Private Cloud.....	78
Measuring the Value of a Cloud .....	80
Changes to Capital Cost.....	80
Changes to Operational Cost.....	80
Labor Costs .....	80
Infrastructure Maintenance .....	81
Facilities Operations.....	81
Simplified Accounting.....	81
Summary .....	82

## **Copyright Statement**

© 2010 Realtime Publishers. All rights reserved. This site contains materials that have been created, developed, or commissioned by, and published with the permission of, Realtime Publishers (the “Materials”) and this site and any such Materials are protected by international copyright and trademark laws.

THE MATERIALS ARE PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE AND NON-INFRINGEMENT. The Materials are subject to change without notice and do not represent a commitment on the part of Realtime Publishers or its web site sponsors. In no event shall Realtime Publishers or its web site sponsors be held liable for technical or editorial errors or omissions contained in the Materials, including without limitation, for any direct, indirect, incidental, special, exemplary or consequential damages whatsoever resulting from the use of any information contained in the Materials.

The Materials (including but not limited to the text, images, audio, and/or video) may not be copied, reproduced, republished, uploaded, posted, transmitted, or distributed in any way, in whole or in part, except that one copy may be downloaded for your personal, non-commercial use on a single computer. In connection with such use, you may not modify or obscure any copyright or other proprietary notice.

The Materials may contain trademarks, services marks and logos that are the property of third parties. You are not permitted to use these trademarks, services marks or logos without prior written consent of such third parties.

Realtime Publishers and the Realtime Publishers logo are registered in the US Patent & Trademark Office. All other product or service names are the property of their respective owners.

If you have any questions about these terms, or if you would like information about licensing materials from Realtime Publishers, please contact us via e-mail at [info@realtimepublishers.com](mailto:info@realtimepublishers.com).

[**Editor's Note:** This book was downloaded from Realtime Nexus—The Digital Library for IT Professionals. All leading technology books from Realtime Publishers can be found at <http://nexus.realtimepublishers.com>.]

## Chapter 4: How Cloud Computing Will Help Your Business

---

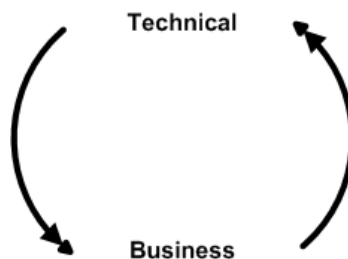
Cloud computing changes the way we do business. Much of the coverage of cloud computing has focused on the technical aspects of this computing model: the consolidation of servers, virtualization, security, and so on. This is understandable, as you must have a clear idea of what cloud computing offers from a technical perspective before you can appreciate what it can do for you from a business perspective.

This chapter turns attention to the business side of cloud computing. In particular, this chapter considers the following:

- How cloud computing can help your business
- Assessing current capabilities
- Introducing cloud computing as a new model for consumption and delivery
- Measuring the value of a cloud

The discussion will start by identifying key business priorities; move on to looking at the current state of IT infrastructure, services, and procedures; describe how to transition those capabilities to a cloud-enabled enterprise; and finally assess the financial value of a cloud to the organization.

Cloud computing has and always will be a rich technical area, but its application to real-world business problems requires the examination of organizational issues that range from bordering on the technical to being the provenance of business executives responsible for overall strategy.



**Figure 4.1: The technical aspects of cloud computing shape what can be done from a business perspective; the business drivers determine how cloud computing is applied.**

## How Cloud Computing Can Help Your Business

Adopting cloud computing is a major change from the traditional distributed systems models many of us use today. No rational business person would make such a fundamental change to core infrastructure without understanding the consequences for the businesses. After all, if your current computing and storage systems are meeting your needs, why change? Why bring on the risks associated with a new technology. Certainly, there is some allure to being on the cutting edge and having the latest technology, but chasing technology trends for their own sake is not a sound strategy for long-term business success. Instead, technology is adopted in service to a business strategy.

Moving your business to the cloud should begin with considerations that have nothing to do with clouds, at least not yet. Cloud computing is a solution; the first question to ask: What is the problem? To get to the answer to that question, you need to:

- Identify business priorities
- Identify operation inefficiency
- Identify barriers to innovation

These steps help you identify what the business is striving to achieve and what is hindering those efforts. Only after you understand that can you turn your attention to addressing the problems that thwart, delay, and increase the cost of your business operations.

Before proceeding to consider priorities, inefficiencies, and barriers to innovation, let's consider a hypothetical scenario that highlights the issues relevant to aligning business goals and technology infrastructure.

### Business and Technology Alignments: The Ideal vs. Reality

Consider a healthcare provider with several hospitals, tens of clinics, and hundreds of doctors serving thousands of patients. As part of a strategic plan to improve the quality of services while controlling costs, executives at the healthcare provider decide to disseminate information on patient conditions, treatments, and outcomes. The executives believe, with sufficient feedback on the results of treatment choices along with details on the cases where particular treatments work and do not work, physicians will be able to reduce uncertainty associated with selecting treatment options.

To implement this plan, the healthcare provider will have to:

- Create a consolidated reporting system such as a data warehouse
- Develop procedures for extracting and loading data from multiple sites into the data warehouse
- Create a reporting infrastructure to deliver information to physicians in a way that is easy to use and fits with their work patterns
- Establish governance over the data warehouse and reporting procedures to ensure compliance with HIPAA and any other relevant regulations
- Define a mechanism to collect feedback from users to improve the system

With the high-level requirements in place, the next step is to determine how the IT department will proceed to implement the plan. Some of the issues that would likely arise include:

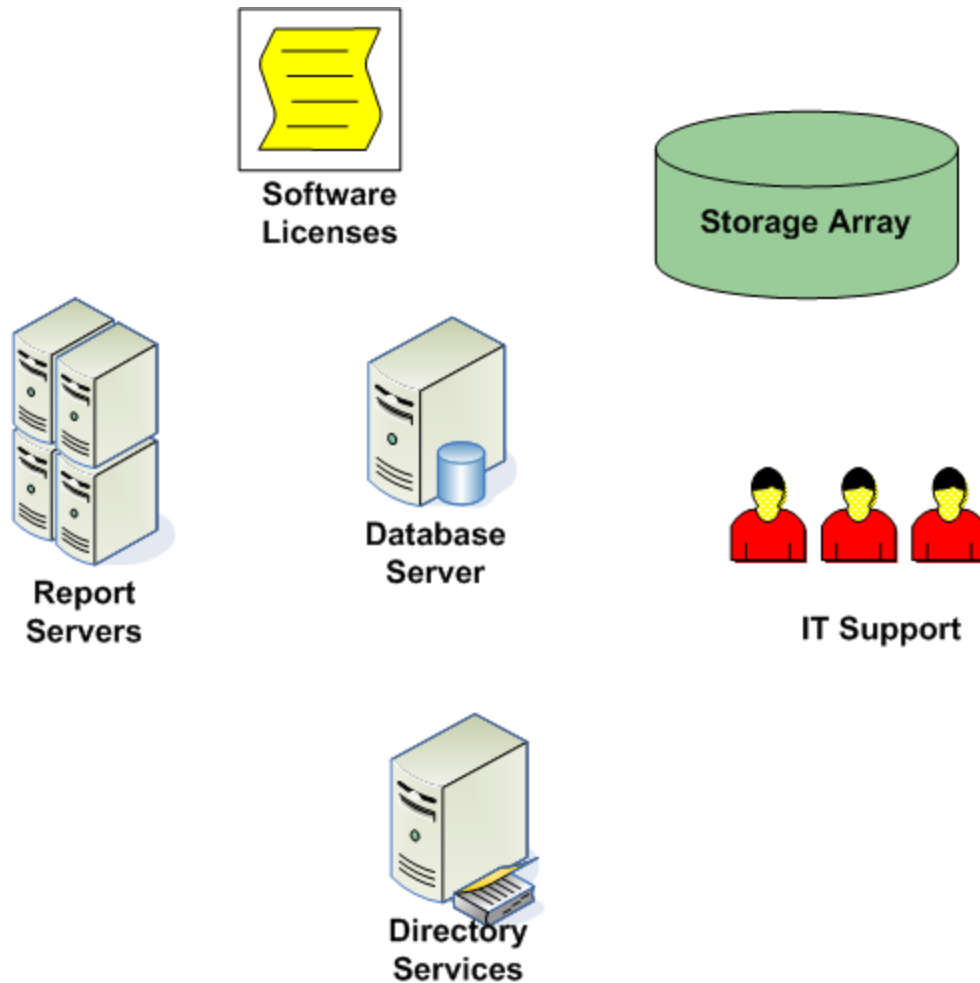
- Acquiring servers and storage to house the data warehouse
- Purchasing licenses for database, reporting, and extraction, transformation, and load (ETL) tools
- Assembling a team to install and configure the infrastructure once it is acquired
- Designing logical and physical data models for the data warehouse
- Developing reports
- Establishing access controls over reports and data
- Creating a support team to monitor data warehouse processes and provide end user support

There are other items you could add to the list, but the list is sufficient to demonstrate the potential drag that IT operations can have on business initiatives. First, though, let's depict an ideal scenario.

IT has sufficient server and storage capacity for the data warehouse. Development work can begin immediately. Fortunately, IT has standardized on a relational database, a data warehousing methodology, and reporting tools. These applications already work with the identity management system in place at the organization, so access controls can be readily established and managed. The support services group within IT is already familiar with these standardized tools, so there is minimal marginal cost to support another set of users.

This ideal scenario is one in which infrastructure, standardized applications, and support services are in place and readily available for new initiatives. To many, this is a fantasy; the reality that many of us have experienced in projects like this is far different from this scenario. Here is a version of the scenario that might ring true for more readers.

The requirements for the reporting project outstrip the available budget. Requirements will have to be prioritized and some features will have to be delayed until later phases. There is insufficient storage available to the business department that owns this project. (There is plenty of storage on another department's disk array, but organizational boundaries rule out using it.) Hardware will have to be procured and installed. Rack space and cabling are a problem that can be worked out with the infrastructure management group, which, given their backlog, will be in a few weeks. The company has a site license for the database software but this project requires several additional packages that will have to be purchased. Several reporting tools are used in other business intelligence projects, so an internal evaluation will be done to determine the best tool for this effort.



**Figure 4.2: Innovative IT projects require an array of technology and services that can be difficult to coordinate and integrate.**

The collective effect of seemingly small and, in many cases, expected problems is to slow down and prolong the implementation. In this scenario, the needs of the business for rapidly deploying an important medical decision support tool is hampered by the way budget for projects, allocated resource across organizational lines, failure to standardized as much as possible on software, and tax support staff with inefficient deployments.

### Identify Business Priorities

One of the most important aspects of successful IT services is that they align with business goals. That is a short way of saying IT services support business objectives in a straightforward manner and do not introduce unnecessary cost, delays, or other burdens on a business strategy.

Although there is no set of priorities that would apply to all or even most companies, common priorities include:

- Controlling costs
- Expanding market share in mature industries
- Expanding into new markets in growth industries
- Improving customer service
- Improving customer retention
- Increasing cross selling

Whatever the priorities and their relative importance, it is critical to identify these for a business. Knowing these will help determine whether and how cloud computing can help your business. For example, if cost control is a top priority, increasing server utilization through virtualization and increasing storage utilization through consolidation with cloud computing architectures can help. If improving customer retention is important, you may need to invest in advanced analytics, such as data mining and statistical analysis, to detect early warning signs of churn. Advanced analytics can be compute-intensive and is a good application for cloud computing. Of course, knowing your business priorities may lead to the conclusion that cloud computing is not something you need at the moment. Whatever your conclusion, if you start with business priorities, you will at least justify why or why not to pursue cloud computing or any other technology.

#### **Caution**

To be clear, cloud computing is not a panacea that will solve all your problems. There are times when cloud computing is not the right solution. It may be an appealing option at a later time, but your business may not be in a position to move to the cloud until it improves its IT governance practices, for example.

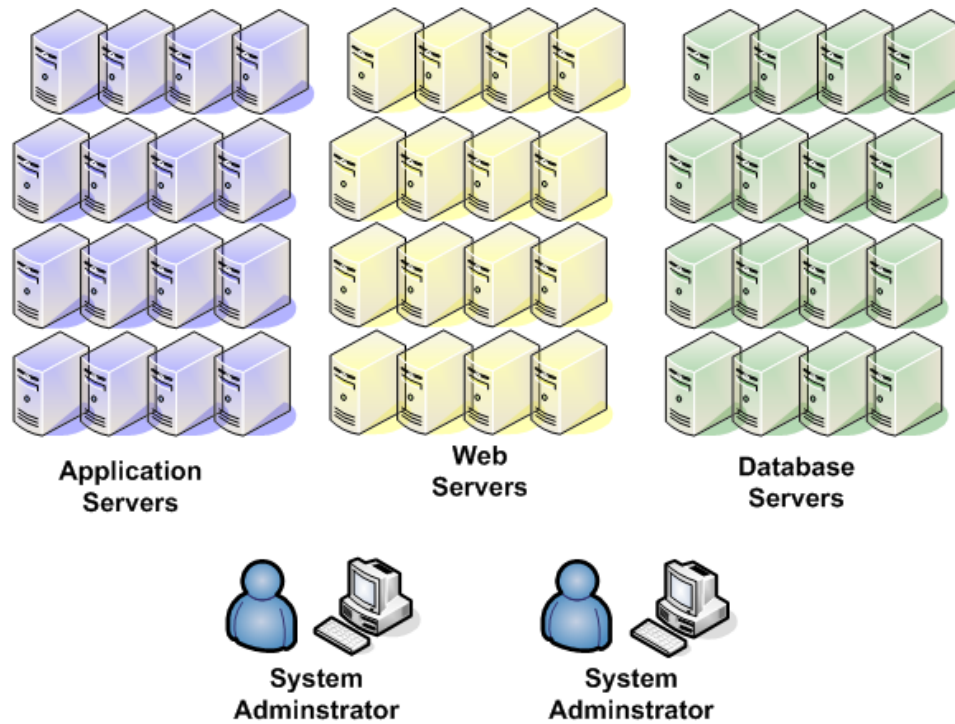
### **Identify Operational Inefficiencies**

Operational inefficiencies are a drain on the bottom line. When an employee has to perform ten steps to complete a task that could be done in six steps, the business loses productivity. When servers are powered on and functioning but not running productive jobs, the business is realizing an opportunity cost as well as incurring unnecessary energy costs. Operational inefficiencies, ironically, are often found in IT departments that have traditionally been a source of increased productivity. Operational inefficiencies in the IT realm come from the way we deploy and utilize hardware and the way we manage software.

Low server utilization is a common inefficiency. Prior to the widespread adoption of virtualization, many organizations used a “one application, one server” approach to deployment. This approach minimized problems with conflicting requirements and allowed administrators to manage servers and applications as a tightly coupled unit. The price we paid for this was wasted CPU cycles.



Less obvious was the management inefficiencies that came with the one application, one server model. Configurations could be tailored to individual applications, so an IT group could soon find that many of its servers had different operating system (OS) components installed, with different applications and services and all requiring slightly different support. As a result, the number of servers a single administrator could manage was limited. Had a standard configuration been developed for a limited set of roles, systems administrators would have less variation to contend with, and this would result in more productive systems management.



**Figure 4.3: Standardizing server configurations is one way to reduce systems management inefficiencies.**

### Identify Barriers to Innovation

The scenario described earlier shows how innovation can stagnate because of technology barriers. It is important to note that the barriers in that scenario were not caused by poor management or unskilled IT professionals; the problem arose from the constraints on procuring new hardware, configuring software, and ordering a sequence of deployment events that account for a range of dependencies between steps.

The scenario can help you see the different types of barriers to innovation that can creep into IT operations:

- Delays in procuring and deploying hardware
- Initial capital costs
- Ongoing operational costs
- Insufficient support staff
- The need to evaluate, select, and coordinate multiple software components

Any or all of these can be significant barriers to innovation. In the time of a globalized economy, customers have more options than ever before, businesses have access to a wider pool of suppliers and business partners, and the list of potential competitors is more likely to grow than not. Add to this list the demands on companies to consistently meet performance expectations quarter after quarter, and you see that barriers to innovation can be a potential long-term drag on the company overall.

The first steps to understanding how cloud computing can help your business is to formulate a clear picture of business priorities, pinpoint operational inefficiencies, and identify barriers to innovation. These three elements comprise the key business drivers that can guide the successful use of cloud computing in your organization. As noted earlier in this chapter, business requirements drive technology-based solutions, but before adapting new technologies, it helps to have a clear understanding of current technical capabilities.

## **Assessing Current Capabilities**

Technology capabilities are a combination of hardware and software infrastructure within an organization as well as the management practices that govern the use of that technology. For the purposes of understanding the role of cloud computing in improving business services, let's consider several types of capabilities:

- Infrastructure
- Platforms
- Applications
- Governance
- Management and reporting

It is important to have a clear and comprehensive understanding of these capabilities because they are all relevant to adopting the cloud computing model. Cloud computing is an evolutionary advance in computer architecture and service management; it improves on what came before it but does not represent a wholesale replacement. Sound management practices, software development life cycle methodologies, and systems administration practices are as relevant to delivering services through a cloud as they are to other delivery methods.

### Infrastructure Capabilities

IT infrastructure for the purpose of this discussion includes server and storage hardware as well as networking components. When assessing current capabilities with regards to infrastructure, consider:

- The inventory of servers currently in place
- The geographic location of servers
- The costs of maintaining each server
- Utilization metrics for servers
- Network infrastructure between sites

At the end of the infrastructure assessment, you should have a clear idea of overall server utilization. If your operations are similar to most, you will have many servers running single applications, and those servers were configured to handle peak, not average, capacity. If this is the case, a move to cloud computing is an opportunity to consolidate servers and decommission those with high maintenance costs, no standard configurations, and relatively low performance. Such consolidation can have an immediate impact on the power and cooling costs of a data center.

There may also be an opportunity to consolidate data centers or at least servers currently located in remote offices. Reducing the number of sites can help ease management overhead and streamline IT operations such as backups.

### Platform Capabilities

Platforms are the OSs and application stacks that run on a company's IT infrastructure. Enterprises typically have a number of platforms:

- Windows
- Linux
- Unix
- Mainframe OSs

Windows and Linux often run on servers that provide specialized functions, such as email servers, content management servers, databases, and directory services. Unix and mainframe OSs are typically found on enterprise-scale computers running high-volume, mission-critical applications.

Cloud infrastructure can be built using low-cost commodity hardware, so such hardware are ideal candidates for hosting Windows and Linux platforms; of course, Unix OSs run on these servers as well.

For the purposes of assessment, you should collect information about:

- The number and version of Windows OSs
- The number, version, and distribution of Linux OSs
- The application stacks that run on these OSs

The goal here is once again to consolidate as much as possible.

### **OS Consolidation**

Standardizing on a reduced number of platforms will reduce systems management tasks and provide a step toward the type of self-service management that is such an important factor in cloud computing's ROI. Standardizing in this case does not mean committing to using only Windows or only Linux but to reducing the amount of variation in the platforms. For example, if a department is still running an instance of Windows Server 2000, this is a good time to move those applications to Windows Server 2008. Similarly, if several distributions of Linux are currently supported, consider reducing that number. It may not be possible to find a Linux distribution that is optimal for all needs, but you might find you can use fewer different distributions than you currently have.

### **Application Stacks**

Application stacks are middleware that reduces dependencies between applications and OSs. When applications are written directly to an OS, they can be difficult to port. Even similar OSs, like different versions of Unix, can harbor enough differences to make porting software difficult. Application stacks and middleware abstract low-level OS details and provide a consistent programmatic interface and set of services. They will be just as important in a cloud environment as they are in today's distributed system environments.

Common application stacks are:

- Microsoft .Net
- LAMP (Linux/Apache/MySQL/Perl or Python)
- J2EE (Java 2 Enterprise Edition)

Application stacks are chosen for their fit with system requirements and the skills of developers working on the applications. Moving applications from one stack to another can be a considerable undertaking, so there is probably less opportunity to consolidate at this platform level. Just as important, though, you will want to ensure that all application stacks currently in use and needed in the future are supported in the cloud.

### *Microsoft .NET Framework*

Microsoft .NET Framework is a development framework for building Web applications for Microsoft platforms. The framework includes several components:

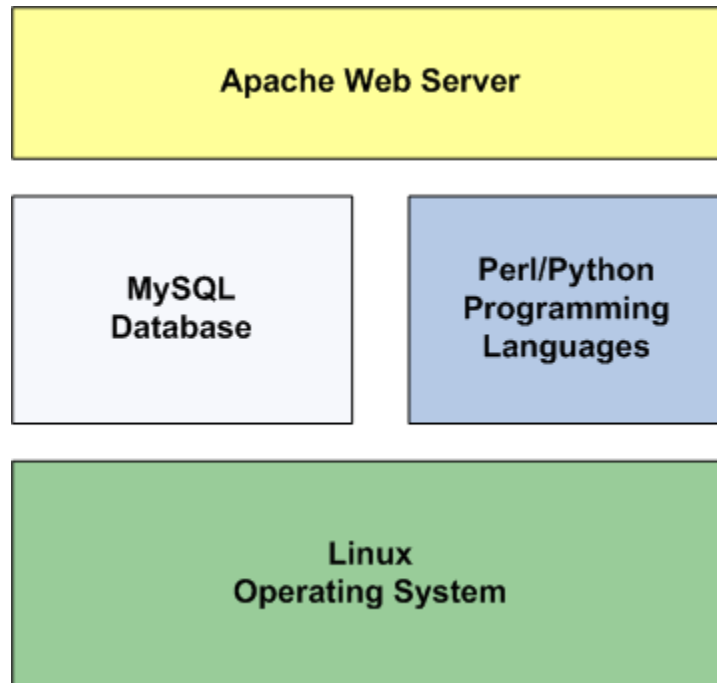
- A common language runtime that acts as an abstraction layer above OS functions
- Base class libraries
- Support for both compiled languages, such as Visual Basic and Visual C#, as well as dynamic languages such as IronRuby and IronPython
- Windows Presentation Foundation, a user interface (UI) framework
- Silverlight, a set of .NET tools for building rich Internet applications (RIAs)
- Windows Communication Foundation (WCF) for service-oriented architectures
- ADO.NET, a set of data access services
- Windows Workflow Foundation

Not surprisingly, the .NET Framework is designed to leverage SQL Server database and other OLE/ODBC data sources.

### *LAMP (Linux/Apache/MySQL/Perl or Python)*

LAMP is a set of commonly used open source systems for building Web applications. Unlike the Microsoft .NET Framework, the individual components of this set of platform tools had long and well-developed histories prior to the advent of LAMP. Each of the four components provides a basic service commonly needed in Web applications:

- Linux is the OS underlying the LAMP stack
- Apache is the Web server and related modules that may be installed as needed for particular applications
- MySQL is a popular open source database suitable for a range of applications sizes and needs
- PHP, Perl, and Python are scripting languages used to implement custom application functions



**Figure 4.4: The LAMP stack consists of a small number of independently developed open source components that are commonly used together for Web application development.**

### *Java Platform Enterprise Edition*

The Java Platform Enterprise Edition, sometimes referred to as J2EE, is a middleware framework designed for deploying distributed Java applications. Like the Microsoft .NET Framework, there are multiple components providing a range of services for application developers. These include:

- Enterprise JavaBeans, a distributed object container
- Java Transactions API
- Java Messaging Service API
- Java EE Connector Architecture
- Java XML Streams
- Java Persistence API
- Java Server Faces, a UI framework

These components are bundled together and incorporated into Java application servers. The application servers sometimes run higher-level additional application development components, such as portals and content management systems.

As you can see from the descriptions of three common application platforms, they provide many of the same functions but do so with fundamentally different tool sets. When assessing existing capabilities, it is important to catalog all the platforms and main platform components used so that they can be supported in the cloud as well.

### Application Capabilities

Enterprises run a wide range of applications and many of these are suitable for running in the cloud. The goal of assessing application capabilities is to determine:

- The relative priority of moving an application to the cloud
- The difficulty to move the application to the cloud
- Changes to application management practices that may be needed after the move to the cloud
- Potential risks and mitigating strategies

When it comes to prioritizing moving applications to the cloud, you should look for those systems that are (a) under-utilizing the servers they run on, (b) are running below needed performance levels because the hardware does not adequately serve current loads, or (c) have peak demands that could take advantage of elastic allocation of CPU and storage capacity of the cloud.

At the same time, you want to avoid immediately moving applications to the cloud that may have special requirements. For example, high-security applications that would require any deleted data be not only deleted but overwritten multiple times to reduce the risk of unauthorized reconstruction of that data. (Deleting data can be done by marking a data block as available for use, so old data can continue to reside on the disk even after files have been logically deleted.) It is not the case that these applications can never be moved to the cloud; they can once security procedures are in place to meet the application requirements.

Running applications in the cloud and on a dedicated server will require different management routines. For example, a cloud storage provider may provide sufficient redundancy in data duplication that you may reduce the number of backups performed. Also, as departments will likely be billed for the time virtual machine instances are running, they will want to optimize their workflows to keep the virtual servers utilized as much as possible when they are running.

Billing rules should also be considered when scheduling jobs. For example, if a department is charged for a full hour of virtual server time regardless of how much of that hour is utilized, it would be best to schedule jobs continuously rather than shutting down and restarting. Of course, this assumes that jobs can be scheduled together that require the same platform. As even this simple scenario shows, the way you manage cloud applications will have to account for new billing structures and server use patterns.

Another area you need to consider during the application assessment is the risk of moving applications to the cloud. These risks include:

- If using a public cloud provider, ensuring security and compliance standards can be met
- Being able to run applications on a standardized platform that may require specialized runtime libraries or utilities
- Having access to virtual machine images that support the application; this can be an issue if a cloud provider only offers the latest patch version of an OS and the application does not run correctly under that particular version

Moving applications to the cloud entails sharing control with the cloud provider. This may be less of an issue when using a private cloud, but it still must be considered at the level of application requirements. These considerations with regard to applications represent just some of the broader issues that must be addressed as part of governance processes.

### Governance Capabilities

A capabilities assessment should include an assessment of governance practices as well. Although cloud architectures are fault tolerant and resilient, the governance practices for clouds are a potential single source of failure. Poor governance affects all users of the cloud.

Governance of cloud operations is required for all types of clouds: private, public, and hybrid. The policies that are implemented will vary by type of cloud, but in general, they will include:

- Complying with government and industry regulations
- Defining and enforcing audit controls and security procedures
- Establishing cost allocation and cost recovery policies
- Setting policies on the management of the service catalog
- Adjusting existing policies to accommodate cloud services

These governance requirements should not be new with cloud computing. The need for governance is independent of IT architecture choices. As noted earlier, though, the cloud changes the way you deliver services and provides new opportunities to change management or governance policies. For example, change control policies may become more flexible with regards to platform-level changes because multiple versions can coexist in the service catalog.

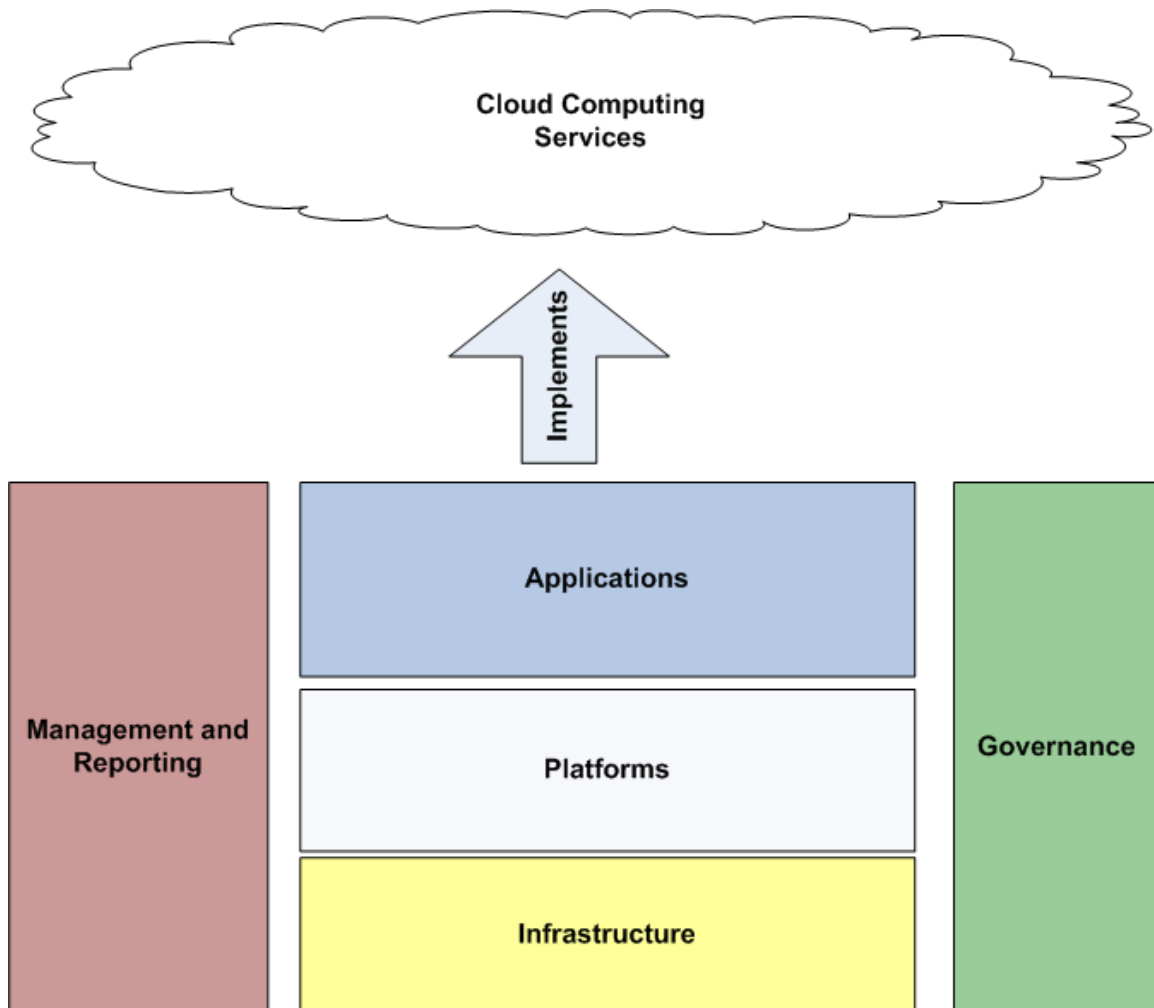
The goal of the governance capability assessment is to understand the mechanisms that are already in place to guide IT operations, identify weaknesses, and make necessary changes. Cloud computing will not improve governance practices, but poor governance can eventually undermine the value of the investment in cloud computing.



### Management and Reporting Capabilities

With cloud computing, service consumers have greater control over how they use computing and storage resources. To optimize their use of these resources, they need information about their workloads, levels of utilization, costs, and other metrics. Management reports are the key to delivering that information.

Reports and data on cloud usage should be available for both front-line managers responsible for scheduling jobs and budgeting for services and for back-office billing operations. Front-line managers should have access to near real-time billing information on CPU utilization and storage allocations so that they can tune workflows. They should also have comparative historical data so that they can detect trends and properly plan for future needs. When a private cloud is used, back-office billing systems will need to accommodate billing or charge backs for cloud services. Existing financial reports would then provide an additional set of reports for front-line managers.



**Figure 4.5: Organizational capabilities in the form of infrastructure, platforms, applications, governance and management, and reporting enable the deployment and use of cloud computing services.**

An assessment of organizational capabilities, spanning existing infrastructure, platform, applications, governance and management, and reporting procedures will provide an organization with a starting point for introducing cloud computing services.

## **Introducing a New Model for Consumption and Delivery**

Introducing cloud computing can be done in two ways: by using a public cloud or by using a private cloud. We will focus most of our attention on the latter, but we will briefly address the use of public clouds.

### **Introducing Public Cloud Consumption Model**

Public clouds can be introduced quickly for small, experimental evaluations that do not involve confidential data, specialized workflows, or complex security requirements. In a very short time, a department-level manager could:

- Establish an account with a public cloud provider
- Upload data for analysis into the public cloud provider's storage system
- Select from the public provider's catalog of OSs and other platform software
- Allocate the necessary number and types of servers
- Run the job
- Shut down the servers, collect the results, and complete the task

This type of isolated, tactical use can also be done in cases where confidential data, specialized workflows, or complex security requirements exist, but it would take significantly more planning, along the lines of what we will be describing shortly in the discussion of a private cloud deployment.

Public clouds allow consumers to experiment with the cloud delivery model without fully committing hardware, software, and management to a full-scale deployment. It is also a viable option for meeting peak demands of jobs that are readily moved to a public cloud. Running significant portions of your business services in the cloud for extended periods can certainly be done but will require the type of attention and planning that one finds with the use of private clouds.

### **Introducing Private Cloud Consumption Model**

There is nothing inherent in cloud computing that requires the cloud be owned and operated by another business. Cloud computing is an architecture and a set of services that enable access resources on demand. The infrastructure and services are managed by the provider and used by service consumers. The provider can be a third party offering a service to the public or an IT division within a company offering computing and storage services to other departments within the company.

A private cloud may appear to lack some of the economic advantages of cloud computing, such as lower management costs and no need for capital expenditures. This may or may not be the case with private clouds; the economic benefit will depend on circumstances within the business providing a private cloud. If the business has a large existing infrastructure with low utilization and high systems management overhead, the company could benefit from redeploying their infrastructure to a private cloud. The number of servers could be reduced because fewer will be needed to meet existing demands. Management overhead could be simplified with cloud management software. In cases where capital expenditures are required, businesses can still benefit from spending less on infrastructure than they would if they did not use a cloud-based approach.

Introducing a private cloud will entail changing procedures and practices; these changes fall into three areas:

- Deploying existing infrastructure in a private cloud
- Enabling application services in a cloud
- Managing a cloud

#### **Deploying Existing Infrastructure in a Private Cloud**

The first step is to establish the hardware infrastructure for running the cloud. Existing hardware may be used for this, but of course, it will require planning to ensure existing services are not disrupted during the transition.

The first step is to identify the servers to use for cloud services. One of the goals of cloud computing is to increase the server utilization, so you would expect to use fewer servers for the same level of demand. If this is the case, older servers with lower overall performance and higher maintenance costs are obvious targets for elimination. Some of the factors to consider when selecting hardware:

- Number of CPUs and cores in the server
- Amount and speed of RAM
- Network interface card throughput
- Cost of maintenance contracts, if any
- Cost of leasing contracts, if any
- Power consumption
- Cooling requirements
- Standardization

Most of these are common sense considerations. The last, standardization, addresses the fact that you should expect hardware failures in the cloud. Actually, you should expect hardware failure in architectural configuration. By standardizing hardware, you reduce the need to maintain multiple types of backup components and streamline troubleshooting procedures.

Storage hardware should be selected for comparable reasons: speed, capacity, throughput, power consumption, cooling, and so on.

Network capacity and throughput should also be considered at the early deployment state. If data centers are being consolidated, additional network capacity may be required. Also, consider the levels of redundancy on the network to ensure services can continue at needed levels if, for example, one Internet access provider is down.

### **Enabling Application Services in the a Cloud**

Application services begin with a service catalog. This is the set of all OSs, middleware, and applications that will run in the cloud. As with deploying hardware, this is an opportunity to standardize on software components. The advantage of standardizing is that there are fewer pieces of software to manage, patch, and configure and that ultimately leads to reduced support costs.

Software services in the catalog should be based on business requirements. There will be needs for different OSs and application stacks, possibly in multiple configurations. For example, to support existing business services, the service catalog may need to include:

- Windows Server 2008 with .NET Framework
- Windows Server 2008 with Java Enterprise Edition framework
- Linux with LAMP framework

In addition to the applications needed in the existing configuration, there will be additional software needed to manage the cloud.

### **Managing a Private Cloud**

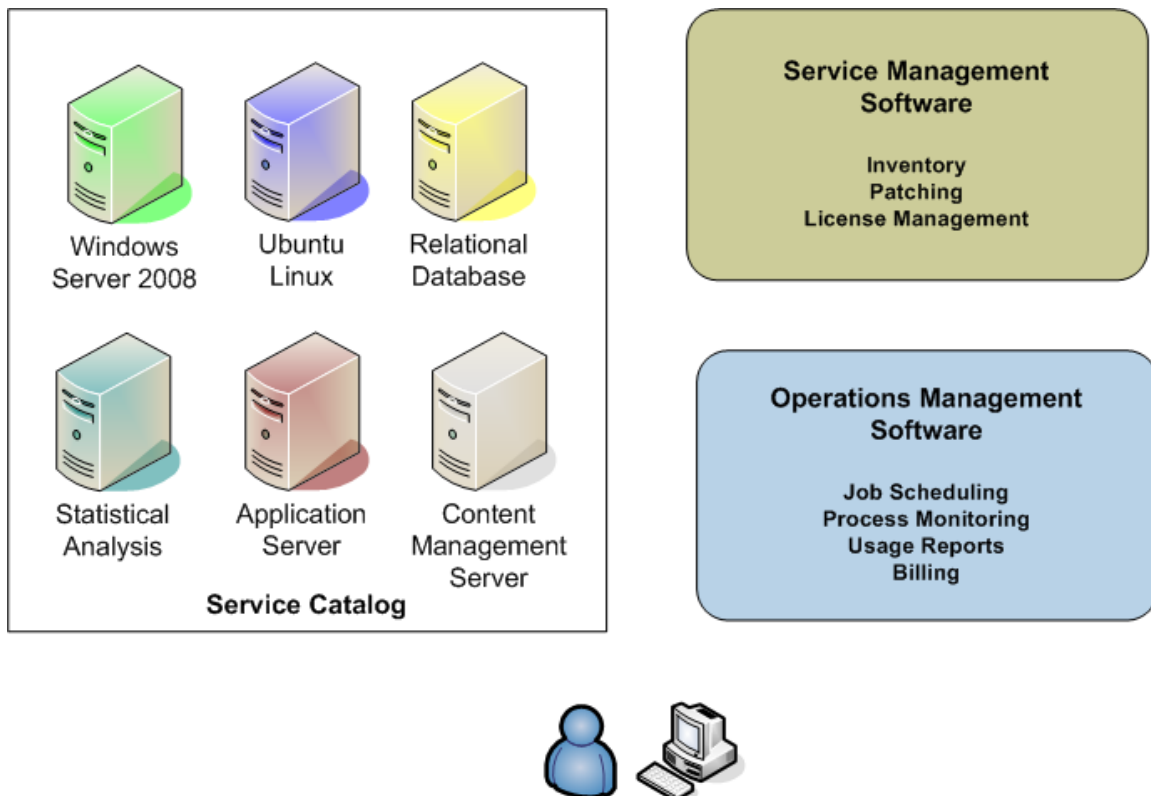
Managing a private cloud requires software and procedures. Operation management software is needed to track the use of compute and storage resources in the cloud. As noted earlier, cloud consumers should have the ability to track their use and costs as they make use of services. They should also have the ability to:

- Monitor their jobs
- Schedule their jobs
- Establish complex workflows
- Track storage use
- Create specialized virtual machine images with custom configurations

Management tools are also needed by IT support staff to maintain the services catalog. For example, systems administrators should be able to track metadata about every item in a service catalog, such as:

- Description of a virtual machine image
- Date it was created
- Applications, libraries, and utilities included
- Patch levels
- Number of times instantiated
- Location of sources used to construct the image

Many of the same service management procedures used in non-cloud environments are still relevant to the cloud. Images will need to be patched, access controls will need to be applied, identities will need to be managed, and charges will have to be made to departments.



Cloud Systems Administration Components

**Figure 4.6: Management components include the service catalog of a platforms and applications available in the cloud as well as management support software.**

## Measuring the Value of a Cloud

Moving to a cloud computing environment will change the IT cost structure and impact both capital cost structures and operational costs.

### Changes to Capital Cost

In a conventional IT model in which departments or service managers use dedicated servers, they often have to plan for capital costs. These are infrequent but significant costs that are budgeted outside the normal operations budget. Although the cost of a server or two may be accommodated in an operational budget, that is not the case for a fully functional application environment.

Consider the costs of developing, testing, deploying, and maintaining applications. For hardware, you would need development and test servers. For small projects, a single high-end server may serve for both as long as each ran in its own virtual environment. The production server may actually be a cluster of servers and a load balancer in order to scale to peak demand. The load balancer will provide some degree of high availability, but disaster recovery procedures dictate a backup set of servers in an offsite location. Storage will be required as well, adding to the capital expenditure. In addition to these hardware costs, there will be the cost of application and OS licenses.

In the cloud, these costs do not go away, but they are reduced. The key is to efficiently share resources rather than dedicate servers and storage arrays to single services or departments. Rather than having multiple service managers develop their own capital budgets, complete with wide ranging contingency funds either explicitly or implicitly added to the budget, central IT can plan for capital costs across a wide base of users. The end result is less capital expenditure because of more efficient use of infrastructure, platforms, and applications.

### Changes to Operational Cost

Cloud computing can prove advantageous for operational costs in four areas:

- Labor
- Infrastructure maintenance
- Facilities operations
- Simplified accounting

### Labor Costs

Labor costs are reduced with the use of self-service management enabled by cloud management software. Consumers of cloud services have the ability to choose the virtual servers they want to run, determine what applications to run, and schedule them when needed. The standardized service catalog reduces the need for costly software configurations, which in turn depend on access to a skilled IT professional to perform the configuration.

The cloud can reduce IT support labor costs in other ways. With a centralized service catalog, updating and patching becomes less labor intensive. For example, if an OS vendor releases a critical patch that has to be pushed to servers, then hundreds of servers may be involved. This requires identifying which servers need the patch, deploying the patch through an automated delivery system, reviewing the results of the patching operation, and manually applying the patch to those servers that failed to be patched correctly using the automated method. This can be a time-consuming burden on IT support staff with other regularly scheduled tasks to complete. The same patch could be applied to images in the service catalog in fewer numbers because only one copy of each configuration is needed. Also, the patch would be available to users of those images the next time they instantiate their virtual machines.

### **Infrastructure Maintenance**

Standardization is a well-established method to reduce costs. Standardizing on infrastructure is no exception. Adding new components, such as servers, to a cloud will have low marginal costs if they are configured similarly to servers already in the cloud. If there are failures (and there will be), the new units are readily swapped in without requiring configuration changes. Inventories of spare components are kept to a minimum as well. Clouds run in a centralized data center, so there is less need for remote office visits to deal with failed hardware.

### **Facilities Operations**

Another contributor to savings in operational costs comes in facilities management. IT infrastructure can consume significant amounts of power leading to high energy costs. Of course, all that power that comes into the data center gets converted to useful computation, but the conversion from electricity to computation is not perfect. The inefficiencies in conversion are realized in the form of heat; heat that has to be removed with costly cooling systems. By driving up the average server utilization, a business can reduce the number of servers needed, which in turn reduces power and cooling costs.

### **Simplified Accounting**

One of the advantages of cloud computing is that it provides a way to standardize computing and storage units of service. For example, a virtual machine running on a dual core processor (or its functional equivalent) for 1 hour can be defined as a unit of computing resource with a standard price attached to it. Similarly, a gigabyte of storage stored for one day could be a unit of storage for accounting purposes. From these fundamental units, you could build pricing schedules that could account for additional costs for OS or application licenses.

With this type of model, cost recovery is simplified. Cloud consumers can readily plan their expenditures. Reporting and integration with financial systems is less complex than if a large number of specialized cases and accompanying business rules have to be accommodated. Cloud computing presents clear cost benefits in both capital and operational costs—as long as proper planning and assessment are done.

## Summary

Cloud computing is an efficient framework for utilizing computing resources. To get the most of your investment, begin by assessing the current state of business and technical operations. This includes identifying business priorities, operational inefficiencies, and barriers to innovation. It also entails assessing the current capabilities in terms of infrastructure, platforms, applications, governance and management, and reporting. Deploying a cloud is a multi-stage process that includes deploying existing infrastructure, enabling application services, and managing the cloud. The value of the cloud will be measured in both capital and operational cost savings.

## Download Additional Books from Realtime Nexus!

Realtime Nexus—The Digital Library provides world-class expert resources that IT professionals depend on to learn about the newest technologies. If you found this book to be informative, we encourage you to download more of our industry-leading technology books and video guides at Realtime Nexus. Please visit <http://nexus.realtimepublishers.com>.