Cloud Computing for Retail

Version 1.0.0

Dec 12, 2009 – Candidate Recommendation

Chairs:

Viswanath Srikanth	IBM Corp
Darrell Sandefur	The Kroger Co.

Members:

Richard Halter	ARTS
Bart McGlothin	Cisco Corp
Sam Palanivel	Lowe's Corp
Ed Johnson	BJ's Wholesale
Johanna Koester	IBM Corp
Leonid Rubakhin	Epicor Corp
Vic Miles	Microsoft Corp
Bill Bittner	BWH Consulting
Graham Hill	Royal Mail
Paul Gay	Epson
Ron Martinez	ProfitBase
Jerry Rightmer	Oracle Corp
David Tran	Verizon Business
Rafael Gonzalez Caloni	Predictix

Contributors:

Martin Wolfe	IBM Corp
Jose Kondooparampil	Sears Canada
David Salisbury	Star Micronics
Sajjan Deogon	Retail Bridge
Martin Folvell	The Norway Group

CONTENTS

Figures

1. TECHNICAL REPORT OVERVIEW

1.1 Introduction

Changes in technology often result in incremental improvements in the retail industry. There are a few examples of innovations that have had significant impact on the cost model of bringing goods to market. Cloud computing has the potential to change all aspects of the retail value chain, resulting in advancements that could create a dramatic shift in the cost model of the modern retail community.

The consolidation of information needed to effectively execute plan, buy, move, and sell activities in retail requires significant computing capability. Cloud computing represents an opportunity for retail organizations to reduce their cost of goods sold in relation to information technology (IT) capital expenditures. At the same time, consumers can have greater access to retail services, often customized to meet their specific needs.

Cloud computing can offer savings in IT resources such as computing, storage, and application services. “The cloud” can provide these services as elastic resources that are suitable for use in existing or new applications, in a wide variety of application types, and by all types and sizes of organizations and companies.

Running applications and storing their data in the cloud can have clear benefits. Rather than buying, installing, and operating its own systems, for example, an organization can rely on a cloud provider to do this for it. Thus, in this scenario, customers would pay only for the computing and storage they use, rather than maintaining a large set of servers just for peak loads.

In a variant configuration, retailers can enjoy significant resource savings by leveraging existing investments in storage, software, and networks by improving utilization levels while applying the same underlying cloud technology, but inside the enterprise.

Also, when written correctly, applications can scale easily, taking advantage of the data centers that cloud providers offer, or the retailers’ own existing infrastructure.

A cloud computing platform can be used in various scenarios. Some examples include the following:

- Retail Price Optimization – A retail provider could choose to provide a pricing engine as a service. This service could rely on input from a client application in a “what-if” scenario or based on a retailer’s stored information to provide consistency across various sales channels (e-commerce, fashion retail, discount outlet).
- Supply Chain Integration – A retailer could decide to develop a software-as-a-service (SaaS) version of an existing on-premises application that receives supply chain signals as inputs to an advance ship notice to stores and distribution centers.

The cloud computing platform lets the retailer focus on its business logic rather than spending time on infrastructure.

- Consumer Loyalty – A retail enterprise can provide basic consumer loyalty services to consumers to allow them to maintain personal preference information, select promotions, or provide feedback.

1.2 Target Audience

This document is intended primarily for IT professionals who plan, design, and implement retail IT infrastructures. The audience for this guide may also include people in the following positions:

- Technical decision makers who determine the appropriate technology solutions for their businesses
- IT infrastructure architects who design systems to provide appropriate service levels to meet the business needs of their organizations
- IT application architects who specify the software and hardware used in a branch store environment
- IT operations managers who manage the services that are provided to the branches
- Consultants and partners who recommend or implement privacy and security best practices to achieve payment card industry (PCI) Data Security Standard (DSS) compliance for their customers
- Independent software vendors (ISVs) that deliver packaged hardware or software solutions to retail customers

1.3 Scope

This technical report presents information about cloud computing as it pertains to retailers and the broader retail community. The document seeks to identify the characteristics of cloud computing that make it compelling for retailers, and attempts to highlight areas in which a cloud-based solution offers strong benefits to retailers.

The report also discusses the key obstacles to adopting cloud-based solutions, including reliability, availability, and security. It concludes with a discussion of issues relating to portability, manageability, and interoperability.

1.4 Out of Scope

Although some consideration has been given to the challenges and benefits of the various deployment models, the report does not provide comprehensive guidance on selection or implementation of the various cloud deployment models (private, public, and hybrid).

The ARTS Technical Committee intends to address further aspects of cloud computing in retail in future revisions of the document. We will look to the ARTS membership community to define the priority and depth of future topics.

This report does not seek to highlight implementations of cloud solutions or the accompanying challenges.

2. WHAT IS CLOUD COMPUTING?

2.1 Defining Cloud Computing

Cloud computing is an emerging computing model by which massively scalable IT-enabled computing capabilities and resources (servers, storage, networks, applications, and services) are delivered as a service to external consumers using Internet technologies. By utilizing both efficient IT service management and a business-centric approach, IT resources and services can be dynamically and rapidly deployed at a significantly lower cost than through traditional mechanisms.

Cloud computing has the following characteristics:

- Virtualization of Infrastructure and Services – Infrastructure and services are made available and accessible virtually over the network. The underlying hardware infrastructure and technology are typically transparent to the consumers.
- Automated Provisioning of Infrastructure Services – Computing capabilities and resources such as servers, storage, networks, applications, and services can be provisioned via a virtualized and standardized interface (i.e., a Web browser) without requiring human interaction with service providers.
- Elastic Scaling – Scalable computing capabilities and resources can be distributed and redistributed dynamically (scale up or scale down) and on demand.
- Increased Accessibility and Availability – Computing capabilities and resources can be accessed by consumers at any time and from anywhere. Plus, additional capacity can be acquired and made available within minutes, without the need to allocate any capital equipment expense.

Furthermore, a cloud computing platform typically offers the capability to support the following:

- Consumption-Based Model – Computing capabilities and resources may be charged to the consumers on a “pay as you go” model and in metered quantity and quality rather than through a one-time license.
- Multi-tenancy – Computing capabilities and resources are shared by several consumers, in the same or different organizations, to take advantage of economies of scale. The security and privacy of each consumer’s data remain intact and transparent to the consumers.

When defining cloud computing for the enterprise retail organization, one must acknowledge the context of the services that will be consumed and how those services are leveraged against the rest of the retail organization’s operational needs. That is to say that cloud computing alone is useful only through the deployment of software services that meet the needs of the retail business.

2.2 Roles within Cloud Computing

The high-level entities involved in cloud computing consist of a service consumer and a service provider, as shown in Figure 1. Service consumers use the services provided through the cloud, and service providers administer the cloud infrastructure itself. It is important to note that open interfaces and open standards should be used in the interactions between these roles.

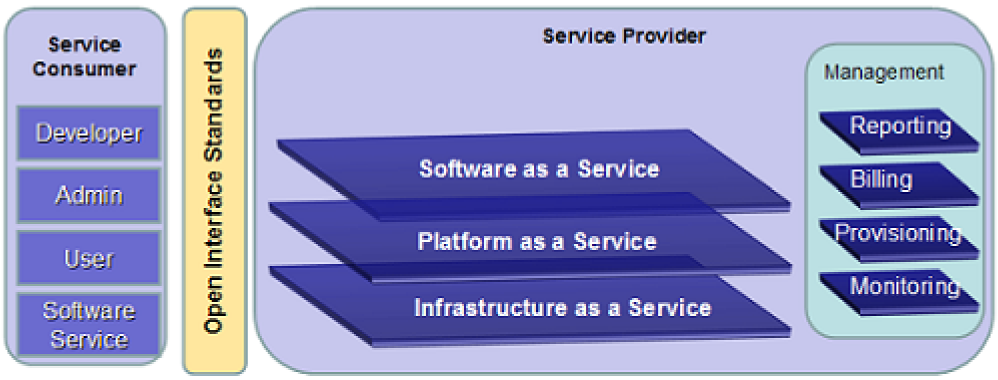


Figure 1: Cloud Computing

The distributed nature of cloud computing allows each entity to perform a role in the model somewhat independent of the other entities. Figure 2 shows the roles within the entities that make up cloud computing. The roles include the following:

- One-to-many service consumers, who consume the services provided
- A service creator, who creates and publishes services to the cloud
- A cloud infrastructure administrator, who manages the services that make up the computing platform

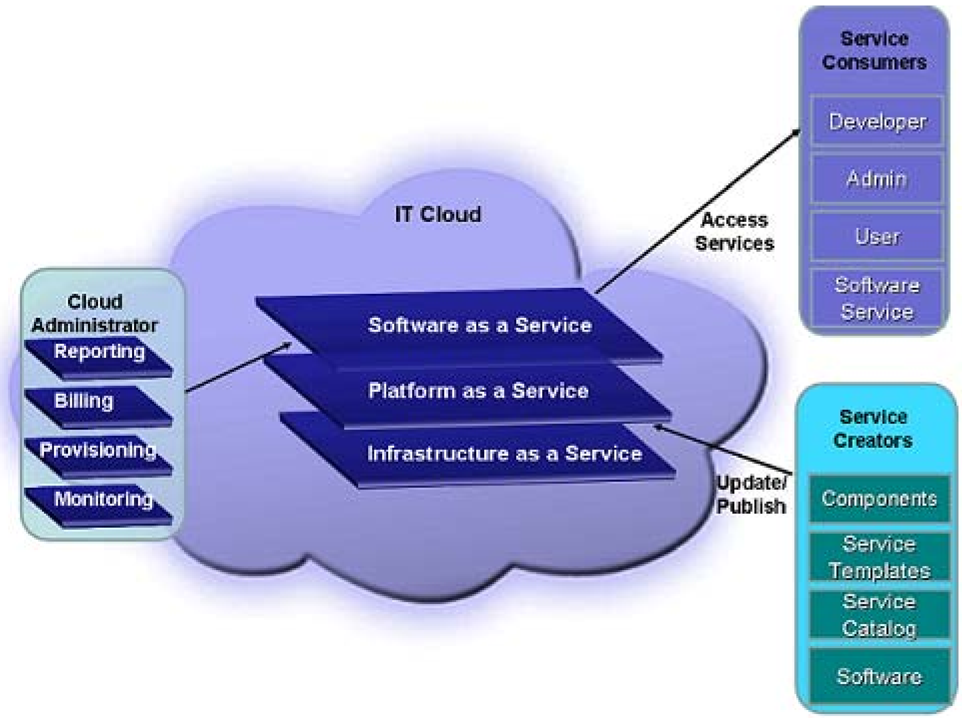


Figure 2: Distributed Cloud Computing

Service Consumer – The service consumer is the end user or enterprise that actually uses the service, whether it is the software, platform, or infrastructure as a service. Depending on the type of service and its role, the consumer works with different user interfaces and programming interfaces. Some user interfaces look like any other application; consumers do not need to know about cloud computing as they use the application. Other user interfaces provide administrative functions such as starting and stopping virtual machines or managing cloud storage. Consumers writing application code use different programming interfaces depending on the application they are writing. Consumers work with service-level agreements (SLAs) and contracts as well. Typically, these are negotiated via human intervention between the consumer and the provider. The expectations of the consumer and the reputation of the provider are a key part of those negotiations.

Service Provider – The service provider delivers the service to the consumer. The actual task of the provider varies depending on the type of service:

- For software as a service (SaaS), the provider installs, manages, and maintains the software. The provider does not necessarily own the physical infrastructure on which the software is running. Regardless, consumers do not have access to the infrastructure; they can access only the applications.

- For platform as a service (PaaS), the provider manages the cloud infrastructure for the platform, typically a framework for a particular type of application. The consumer's application cannot access the infrastructure underneath the platform.

- For infrastructure as a service (IaaS), the provider maintains the storage, database, message queue or other middleware, or the hosting environment for virtual machines. Consumers use that service as if it were a disk drive, database, message queue, or machine, but they cannot access the infrastructure that hosts it.

Cloud Administrator – Crucial to the service provider's operations is the administration layer. At a low level, cloud administration requires metering to determine who uses the services and to what extent, provisioning to determine how resources are allocated to consumers, and monitoring to track the status of the system and its resources.

At a higher level, administration involves billing to recover costs, capacity planning to ensure that consumer demands will be met, service level agreement (SLA) management to ensure that the terms of service agreed to by the provider and consumer are adhered to, and reporting for administrators.

Security applies to all aspects of the service provider's operations. (The many levels of security requirements are beyond the scope of this paper.) Open standards apply to the provider's operations as well. A well-rounded set of standards simplifies operations within the provider's environment and interoperability with other providers.

2.3 Services Offered by the Cloud

The options for obtaining value by leveraging cloud services can be understood through an examination of the various types of services that can be delivered from the cloud. As mentioned previously, services delivered from a cloud typically fall into one of the following three main categories, as depicted in Figure 3 below:

- Software as a Service (SaaS) – SaaS is a software distribution model in which software is hosted by a provider in a central and remote location and made available to consumers, typically over the Internet. SaaS is based on a pay-as-you-go pricing model, which gives consumers the ability to decrease or increase the number of software licenses based on need without having to install or maintain software or hardware or incur ongoing maintenance costs. With SaaS, business applications such as accounts payable and customer loyalty are made available in a virtualized fashion to the end user.

- Platform as a Service (PaaS) – PaaS is a form of cloud computing in which the computing platform (hardware and software) is delivered as a service, typically over the Internet. PaaS enables developers to create, test, deploy, and host applications (i.e., Web applications) quickly without having to bear the cost and complexity of buying and managing the underlying software and hardware. PaaS can also be referred to as "cloudware." With PaaS, Web services, Web 2.0 capabilities, and middleware can all be offered as a platform on top of which applications can be built, assembled, and run.

- Infrastructure as a Service (IaaS) – IaaS is similar to SaaS in that a product is offered via a virtualized interface, typically the Internet, to consumers as an on-demand service; but instead of software, IaaS delivers hardware components such as servers, network equipment, memory, CPUs, and disk space. With IaaS, retail organizations can run their entire operations without having to install and maintain in-house data centers.

The approach to the delivery of these services differs between cloud providers. In order to provide scale, the platform and infrastructure layers are optimized and streamlined in terms of functionality. For instance, when a cloud-enabled application is written and/or deployed, that application may not have traditional relational SQL access to data. This abstraction is what allows infrastructure providers to offer scale across the three layers.

In summary, the cloud infrastructure focuses squarely on efficient utilization of the base infrastructure. This includes virtualization, routing, and storage management. The cloud platform manages the services running on the infrastructure. Finally, the services being provided on or by the cloud are what the ultimate consumers use.

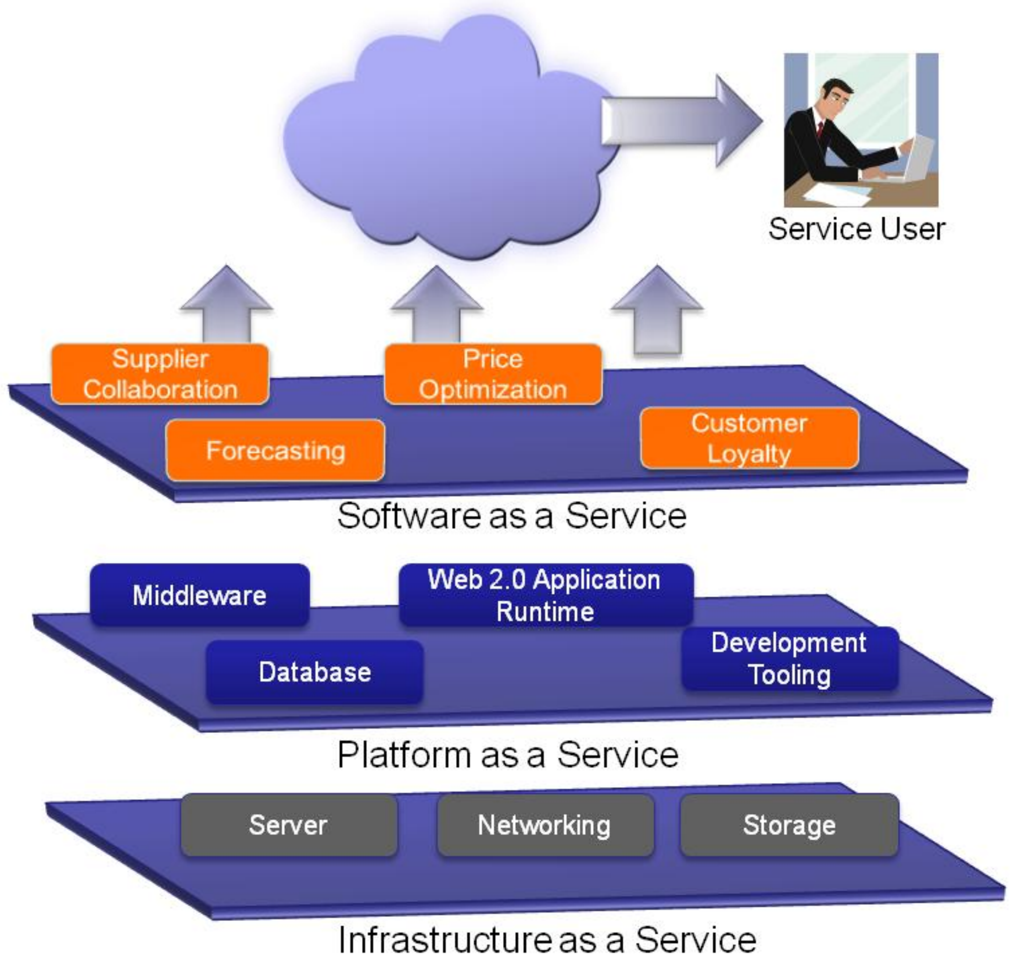


Figure 3: Cloud Services

2.4 What Is Different about Cloud Computing?

Before delving deeper into the types of cloud computing, it is worth examining why cloud computing is attracting so much attention. Cloud computing has evolved on a continuum that includes grid computing, utility computing, and application service provisioning (previously referred to as SaaS) as well as advances in efficient utilization of infrastructure resources.

Cloud computing thus has two dimensions along which it offers incremental benefits:

- Improved service delivery
    - Ability to bring service capability on- and offline quickly
    - Combining service capabilities more easily (sometimes known as mashing up services)
    - Wide accessibility of services on a dynamic scale from moderate to massive
- Efficient utilization of infrastructure
    - Software
    - Servers
    - Storage
    - Network

2.5 Types of Clouds

In the general industry, two forms of access to a cloud computing model are beginning to develop: private cloud and public cloud, with a third model that straddles the two, called hybrid cloud.

Private Cloud – This form of cloud computing is an enhancement of traditional enterprise data centers that have always been a part of an organization's architecture and provisioning. The infrastructure within a private cloud is typically operated solely for a user organization and may be owned and managed by the user organization or hosted by a third party and may exist on-premises or off-premises. It provides restricted access to the computing capabilities and resources to be shared among employees, internal departments within the user organization (human resources, IT, marketing, etc.), and external partners (distributors, manufacturers, etc.). Private cloud computing helps drive efficiency, standardization, and best practices in the services it provides while retaining greater customization and control than public clouds would permit.

Public Cloud – The cloud infrastructure within a public cloud is owned and managed by an organization selling cloud services and is made available to the general public or a large industry group. In this model, computing capabilities and services such as standardized business processes, applications, and/or infrastructure services are accessed by multiple subscribing clients on flexible pay-per-use basis.

Hybrid Cloud – The infrastructure within a hybrid cloud consists of a combination of both private cloud and public cloud features. In this model, computing capabilities and resources are owned and maintained by both the user organization and the cloud provider. Typically, a user organization uses public cloud computing capabilities and services for general and non-mission-critical computing but stores customer and mission-critical and sensitive data within its own private cloud to ensure security.

To compare the benefits of public clouds against those of traditional dedicated services, it is important to evaluate the software services provided by application service providers (ASPs) compared with the software services resident on the public cloud. Similarly, for private clouds, the focus should be to compare the traditional hosted enterprise IT infrastructure available through data centers with the services that are available using a private cloud.

2.5.1 Comparing Public Cloud Services with Traditional ASPs

This section compares the attributes of service delivery from a traditional ASP and a public cloud provider.

The following attributes are typical of a traditional ASP:

- The application is hosted and managed by the vendor on its IT infrastructure.
- Payment is on a subscription or per-use basis.
- Limited front-end customization is offered (for a fee) on a per-client basis.
- An increase in computing capacity is offered but requires multiday turnaround.
- New vendor-specific applications may be offered as a service over time.
- The ASP does not offer any way to combine and integrate various services.

Service on a public cloud possesses the following attributes:

- The software service provider on the cloud (application vendor) and the cloud infrastructure provider (host) are usually distinct entities.
- Cloud computing typically supports a large variety of cloud services, including a range of operating systems, hardware, and software applications.
- Cloud services can be brought on- and offline quickly (turnaround in minutes, not days), permitting rapid provisioning of services.

- Computing capacity can be increased or decreased rapidly (elastic scaling), thus permitting responsiveness to unpredictable demand.
- New services can be built on top of existing services through integration, thus permitting new solutions to be offered as a service.
- Cloud computing offers a flexible pricing model.

Table 1 compares the features associated with delivering software services via a traditional ASP model and a cloud-based SaaS model.

Table 1: Comparing Application Service Providers with Public Cloud Service Providers

Feature	Application Service Provider (ASP)	Software as a Service (SaaS) through a Public Cloud
Hosting and management	The application vendor hosts and manages the application.	The software service and cloud infrastructure providers are distinct entities.
Payment model	Flexible pricing options—per use, subscription—are available.	Flexible pricing options—per use, subscription, and potentially options based on infrastructure usage (storage, network, CPU)—are available.
Number of services / service providers	Typically only vendor-specific applications are available from that ASP.	A wide variety of services from a large number of vendors are easily supported on a cloud.
Increasing/decreasing computing capacity	The computing capacity offered can be increased, but turnaround is usually in days (or weeks).	Computing capacity increases/decreases on demand, in response to changing load on the services.
Bringing services online/offline	Services are usually kept on, as bringing services on- and offline is resource intensive.	Services can be efficiently brought on- and offline, thus providing self-service options and permitting efficient reuse of resources.
Integrating services / building solutions on third-party systems	The focus is on providing a service, not on integrating services at the hosting site.	The cloud permits the combining of services in the cloud, leveraging middleware services in the cloud, and potentially laying out entire architectures and solutions in the cloud.

According to the above comparison, the benefits of the cloud as a hosting environment for software services include the following:

- A large of number of application providers can expose their applications as a service, thus increasing choice for retailers and the retail community. The cloud platform tends to support multiple operating systems and applications written in multiple languages, further increasing choice and encouraging competition.

- Significant support for middleware services, including application servers, queues, and databases, has become increasingly available on public clouds, permitting solution integrators to connect services and build solutions in the cloud and offer them via a flexible payment model.

- Efficient utilization of resources, including the ability to match resource use against demand, is further driving down costs for users of cloud services, who no longer need to plan for peak demand up front from an IT resource provisioning perspective.

2.5.2 Stacking Up Private Clouds against Traditional Data Centers

The attractiveness of services hosted on a public cloud is understandable given the benefits and challenges highlighted in the previous section. However, the infrastructure on which the cloud is built also offers some efficiency benefits that can reduce costs for those retailers who have relied on traditional data centers and other forms of hardware, software, storage, and networking investment to run and manage their enterprise IT.

Traditional IT investment involves the following:

- Software and hardware particular to the needs of the application and its anticipated usage must be purchased, which frequently involves erring on the side of anticipated peak usage. However, if the application is used continuously and is consistently in high demand, a dedicated infrastructure would prove to be a worthwhile investment.

- Each new application purchase typically results in the creation of another stack of hardware and storage, with very limited ability to reuse existing investments in infrastructure. The usage of the application (frequent, rare, or occasional) and predictability of demand determine whether it is worth moving away from the traditional IT infrastructure to a more efficient infrastructure.

- Retailers that maintain their own test and development environments make significant investments in infrastructure that periodically becomes underutilized, as it is too expensive to tear down and build up these environments.

The private cloud environment offers a few advantages for certain types of workloads when compared with traditional data centers and IT environments:

- Efficient utilization of hardware, storage, and networking through the use of virtualization results in a lower overall investment in infrastructure. This is particularly useful for workloads that have unpredictable demand and that are not in consistent heavy use. Running those applications on a consolidated infrastructure can add significant value for the retailer.

- The ability to bring services online and offline quickly (automated provisioning) affords a positive user experience while improving server utilization by keeping only those services online that are in use. This quality is important if services are to be utilized efficiently. This not only reduces hardware, storage, and networking costs but also saves on the number of licenses required for an application.

- The cloud offers the ability to increase computing capacity for any given task automatically, without manual intervention. Capacity is only limited by the physical number of servers, and the greater flexibility of cloud-based services is an obvious benefit for handling unpredictable demand.

- Centralized management of software and IT infrastructure may reduce management costs, but this is directly dependent on the current economics of IT management for the retailer.

Private clouds are also finding traction because they are housed within the enterprise, thus significantly reducing security, availability, and accountability concerns.

2.6 SOA and Cloud Computing

Service-oriented architecture (SOA) is defined in the ARTS SOA Blueprint for Retail as follows: “SOA is a business-centric approach to delivering information technology (IT) capabilities by using common services to perform basic business functions. These services are software modules that fulfill business needs while hiding implementation complexity and increasing maintainability and reuse.”

In other words, the core focus of SOA is to deliver software services to support retail business processes and functions while hiding implementation complexity. To deliver SOA one needs to think about the following:

- The business and business services
- The provision of these business services through IT services and the corresponding architectural layout
- The infrastructure to support the IT services and architecture
- The runtime and delivery environment needed to make these services available, accessible, reusable, and secure

The focus of SOA is on services, and cloud computing offers a platform to deliver the services needed as part of a SOA.

As highlighted above, cloud computing offers significant enhancements to and efficiencies in the infrastructure upon which software services run, offering the availability of services on an as-needed basis, offering the automatic scalability of computing capacity, and permitting services to be accessed via a flexible, pay-per-use model.

Thus, from a SOA perspective, the cloud becomes another source of IT services that the retail enterprise could access and combine with other services within the enterprise to deliver the business functions it needs. The cloud itself may be private or public, but from a SOA perspective, it remains a source of useful services that could be combined and recombined as part of a flexible, agile enterprise architecture.

When the service architecture from the ARTS SOA Blueprint is expanded to include services from the cloud, it appears as shown in Figure 4 below. Note that this diagram is excerpted for illustrative purposes and is not complete. Refer to the ARTS SOA Blueprint for the full diagram.

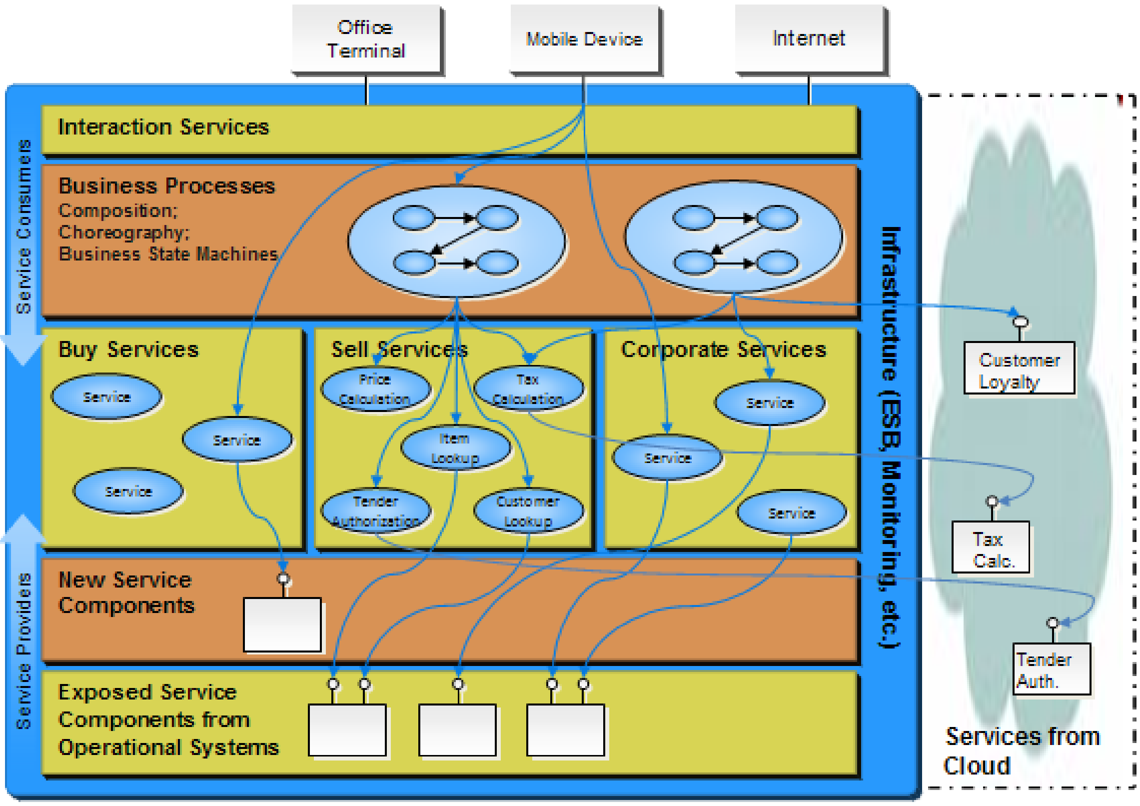


Figure 4: Retail service Architecture with Selected Services in the Cloud

SOA has typically dealt with software services and the ability of the architecture to flexibly integrate these services in response to changing business requirements.

As indicated earlier in this chapter, providing SaaS is a key aspect of cloud computing and cloud services. However, cloud computing extends this capability into the infrastructure and platform as well.

In PaaS, middleware capabilities—including databases, Web-service connectivity, queue services, enterprise service bus (ESB) functions, and more, typically useful for creating a flexible SOA architecture—are also available as services. Figure 4 above shows some of the middleware capabilities needed for a SOA implementation to be available as a service. In order to take advantage of the cloud computing PaaS capability, the retail IT organization should have an advanced level of service integration maturity. The level of service integration maturity has been defined in a model published by the Open Group, the Open Group Service Integration Maturity Model (OSIMM).

OSIMM defines a set of dimensions representing different views (e.g., business, architectural) of an organization as follows:

- Business
- Organization and governance
- Method
- Application

- Architecture
- Information
- Infrastructure and management

Specifically, the OSIMM defines seven levels of maturity, as indicated in Figure 5 below. The level of the architecture dimension needed to effectively implement applications in a PaaS environment is level 5, which is characterized by being SOA enabled.

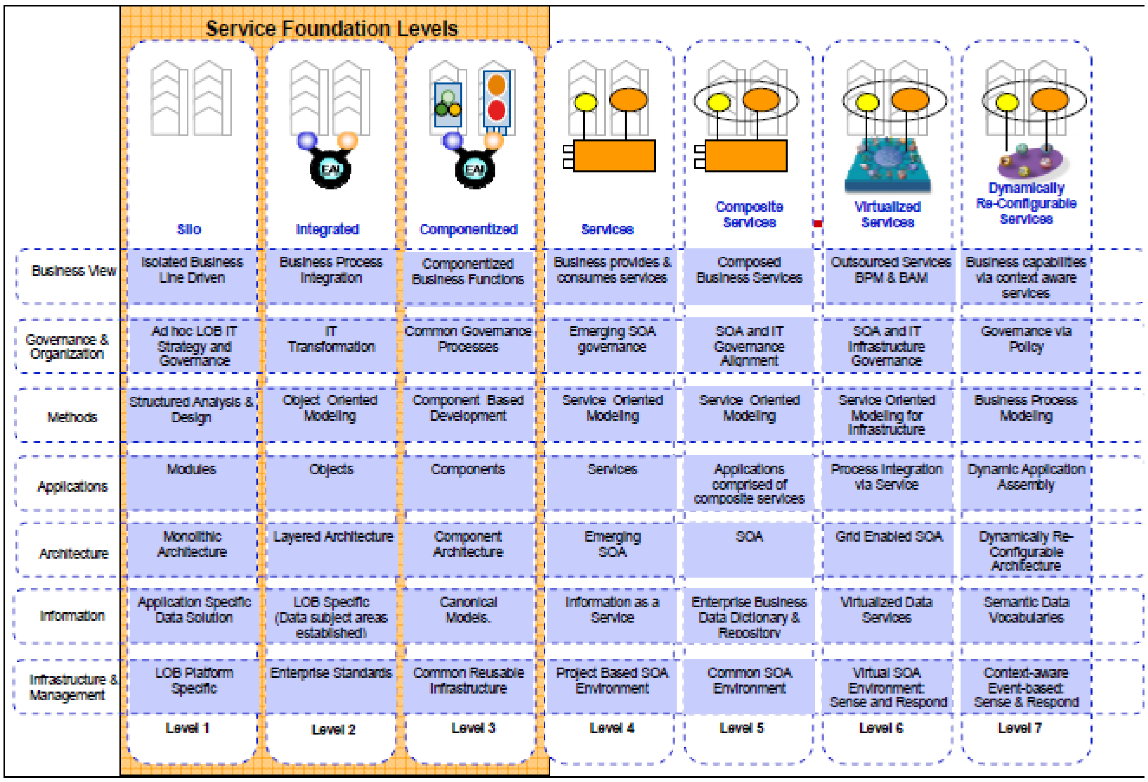


Figure 5: OSIMM Maturity Matrix

The following sections delve deeper into areas in retail where cloud computing can be applied and address considerations in leveraging the cloud for accessing services.

3. APPLYING CLOUD COMPUTING IN RETAIL

Cloud computing can offer significant benefits to retailers if it is adopted in the right areas. The following section highlights some key characteristics of these areas, and the rest of the sections of this chapter highlight specific examples based on the identified characteristics.

3.1 Characteristics of Services in Retail That May Benefit from a Cloud Platform

A few common threads flow through the types of retail services that may benefit from being offered on a cloud platform. These include services that

- face unpredictable demand;
- have a multiparty user ecosystem, requiring collaboration;
- will face increasing consumption over time as additional service consumers (retail customers, business partners, other retailers) are brought online;
- help in the consolidation of software, hardware, and management (instead of requiring applications on multiple desktops or laptops);
- a retailer does not want to develop internal expertise to execute (e.g., geo-specific tax services and human resources)
- permit efficient reuse of hardware investments (through easy provisioning and de-provisioning of software configurations or environments, as frequently seen in test and development environments);
- are computationally intensive (the service is used specifically for short bursts—e.g., price optimization, demand forecasting, and replenishment); or
- need to be migrated easily between public and private clouds as needed (seasonal capacity—moving a service to a public cloud at some time during the year).

Additionally, it is important to highlight the possibility of mashing up services (i.e., combining multiple cloud-based applications) to provide higher-value-added services for retailers; this is a benefit that should be considered when deciding on the many ways to leverage the cloud platform and services available.

Services that are noncritical may form the first set of candidate services that a retailer may seek to leverage on the cloud; as confidence improves, some of the more critical services can be moved to the cloud platform.

When looking at the above from a pure private cloud perspective, the retailer becomes both a provider of the service (even if the retailer hosts a third-party application to provide the service) and a consumer of the service. The consumer perspective remains largely the same, though some of the security concerns do ease. As a provider, the retailer would have to worry about provisioning for peak capacity as well as consistent availability of services, much as with IT infrastructure at the store and enterprise today. However, private clouds do offer efficient use of infrastructure and centralized management, which provide additional value.

Usually, when trying to compare the characteristics of a service suitable for private clouds versus a typical enterprise data center, it is useful to look at the workload characteristics of the service. In other words, if the workload characteristics clearly show the benefits of operating on a platform that is virtualized, is self-provisioned, and has elastic computing capacity, one can safely and profitably move those workloads to a private cloud. Examples include test and development environments, virtual self-provisioned desktops, and more.

3.2 Logical Categories of Retail Services and Cloud Computing

The types of IT services offered from the cloud can belong to any one or more of the categories of services identified in the ARTS SOA Blueprint for Retail reference picture: Plan, Buy, Sell, Move, and Corporate.

These services are hosted on a virtualized infrastructure that hides the hardware, storage, network, and operating system complexity from the service itself. Figure 6 depicts this relationship.

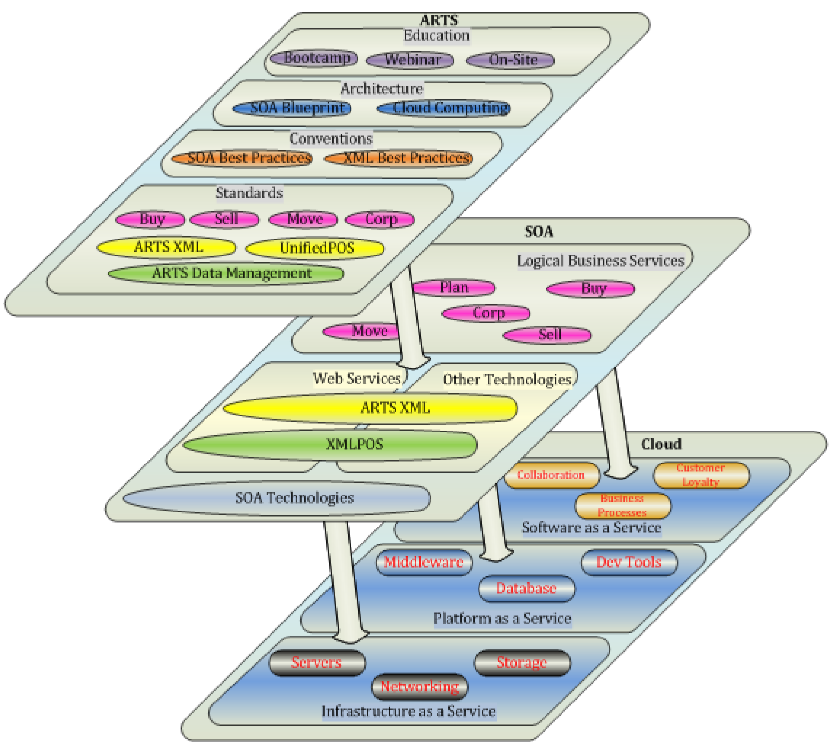


Figure 6: SOA Platform and Services on a Cloud Infrastructure

3.3 Illustrative Set of Examples for Applying Cloud Computing in Retail

This section provides a set of examples that display some of the characteristics identified earlier and illustrate where a cloud-based solution lends considerable value.

3.3.1 Providing Services to Associates in the Store

Challenge: Store associates frequently need access to a range of applications, including access to benefits information, a paycheck calendar, store training, punch-in/out services, inventory services, and more. These applications are provided via a variety of deployment patterns—ranging from hosting on store associate desktops or laptops to in-store servers. In some situations, a store associate may have three or four computers assigned to him or her, each with the various necessary programs resident. This makes the management (operational, upgrades) cumbersome and expensive. Moreover, many of these applications and hardware are sitting idle for long periods of time, unused.

Why Is a Cloud Solution Suitable? A key benefit of a cloud infrastructure is that one can consolidate the servers and software on a virtualized platform, provision services on an as-needed basis, and add computing capacity on demand.

Type of Cloud: Private (typically) or public (as services become available)

Type of Service: SaaS and IaaS

Solution Brief: In this situation, all the applications needed by the store associate can be managed in a centralized cloud platform, consolidating the underlying hardware and software required and providing only thin-client access to the store associates. Store associates can log on and configure any form of "virtual desktop" they want at any given time, or access any particular applications that they are interested in.

- Improved efficiency of hardware and software usage
- Reduced IT management (centralized updates and management of IT)
- Leaner IT in stores

Risks:

- Change in usage patterns
- Disruption to business activity during the technology handover, including the cost of retraining associates
- Regression test and services needed to ensure that nothing is broken after the switchover (preproduction test)
- Need to ensure connectivity to the cloud during operations
- Availability of all required software to run in a cloud environment

3.3.2 Providing Services to Customers in the Store

Challenge: Store owners seek to offer personalized or brand-aware store experiences to their customers. One way to do this is by offering specialized, brand-specific experiences to customers in a store through their mobile devices (after the customers have opted in). The goal is to give customers details about the location of specials, discounts, specific categories of items, and more, making the experience specific to the store. Other similar services, such as self-service solutions, can be similarly hosted on a cloud and delivered to the customer in the store.

Why Is a Cloud Solution Suitable? Handling variable or unpredictable demand patterns is one of the key strengths of cloud computing. Offering services (including self-service options) to thin clients such as mobile devices over public or mobile networks is a natural fit for a cloud solution.

Type of Cloud: Typically private (brand-specific or store-specific customization of service required)

Type of Service: SaaS and IaaS

Solution Brief: Hosting mobile services (or self-service solutions) on a cloud platform allows the store to manage a large number of customers and handle unpredictable demand seamlessly, while also sending messages over the public or mobile network. The retailer-specific mobile application is triggered by an acknowledged presence of the customer in the store (loyalty-card swipe, facial recognition, customer-initiated mobile login, or other means); the application then interacts with the customer's mobile device and sends needed or relevant information as the customer browses through the store. The mobile application can handle one or multiple stores, depending on the configuration. Similarly, other self-service solutions can be hosted on the cloud platform and interact with the user through thin clients.

Benefits:

a) Easy handling of unpredictable demand
b) Leveraging of public or mobile networks for connectivity
c) Customization of solutions on a per-store or store-cluster basis if using private cloud platform

Risks:

a) Connectivity and availability of cloud services
b) Potential limitation of customization options if using public cloud, as most service providers seek to maximize the clients they can serve with the same basic service

3.3.3 Partner Collaboration Solutions

Challenge: There is a growing need for collaboration between retailers, suppliers, and business partners to deliver increased value and service to consumers and to gain a greater percentage of their wallet-share. The need to share customer-related information (buying patterns, loyalty information) and other sales-related information continues to increase.

The nature of these relationships, however, requires the ability to host these solutions and services in an easily accessible fashion; to permit growth in computing capacity as an increasing number of partners, manufacturers, and consumers come on board; and to continue to handle variable or unpredictable daily demand easily.

Why Is a Cloud Solution Suitable? Steady increases in computing capacity, easy access to services for multiple constituents, and the ability to handle variable demand are all qualities that make these solutions an easy target for the cloud computing platform.

Type of Cloud: Public or private clouds will work for these solutions.

Type of Service: SaaS, PaaS, and IaaS. Creation of the solution may leverage middleware services, database services, and other platform services on the cloud in addition to the final application service and the core infrastructure service.

Solution Brief: As one example of a collaboration solution, if a retailer permits the sharing of customer loyalty information with manufacturers or other business partners, they in turn can offer discounts to the customers directly or indirectly. Customers can log on to the loyalty service and choose to load these offered discounts onto their digital profile and redeem them at checkout at the retail store or other business partners' stores.

Hosting such a loyalty service on a cloud platform allows the retailer to efficiently handle unpredictable demand, but it also allows the service to deal with incremental growth as more business partners, manufacturers, and customers come on board and begin leveraging the service.

Other similar forms of collaboration can occur when retailers and manufacturers share daily sales data on an ongoing basis, or when suppliers are seeking to access retail inventory data, and so forth. The cloud platform offers an efficient means of hosting and delivering these services.

Other examples of partner collaboration solutions can be found in merchandising, assortment planning, category management, and pricing. For example, an apparel brand can collaborate with its retail partners to plan floor sets across the retailer's locations. A grocer might also collaborate with its vendors on category management. A wholesaler can collect promotional plans from its retail clients to more accurately forecast sales and avoid stockouts by replenishing more efficiently and effectively.

Benefits:

a) Computing capacity on demand

b) Easy access to services by multiple stakeholders

c) Efficient infrastructure usage

d) Leveraging of middleware and/or database services on the cloud

Risks:

a) Connectivity to the cloud

b) Data security for data living on the cloud (if using public cloud)

c) All the chief concerns with a cloud offering, as highlighted in Chapter 6 on reliability

3.3.4 Analytics on the Cloud

Challenge: Business analytics are a foundational component of retailing. Metrics help retailers make data-driven decisions involving marketing, merchandising, and purchasing. Analytics are also increasingly needed for transactions between business partners such as manufacturers and retailers. Analytics need to be plugged into retail operations and supply chain systems so reporting services are readily available. A key requirement for retailers is to have sufficient computing capacity available so additional partners can be added.

Why Is a Cloud Solution Suitable? Easy access to analytics services on the cloud platform, combined with the ability to adjust computing capacity on demand, including adding incremental computing capacity, makes a multi-partner analytics solution a good fit for the cloud. Data can grow exponentially, depending on which metrics are gathered and for how long. Business requirements can push reporting customization beyond the capabilities of internal development staff.

Type of Cloud: Public or private

Type of Service: SaaS and IaaS

Solution Brief: As described in Partner Collaboration Solutions above, the need for greater sharing of information across the retail ecosystem is increasing. At the same time, analytics needs to be applied on data coming from sources beyond one's own organization (tracking and analyzing Twitter mentions of the brand name, tracking and analyzing fresh-item movement from multiple suppliers and third-party logistics [3PL] companies, and more). Analytics services available on the cloud platform can be tied into one or more of these data streams as needed to obtain the necessary business intelligence (assuming the service can handle the required type of analytics). Over time, analytics on the cloud may develop into a series of standardized services that retailers can exploit profitably without having to invest considerable capital in the space.

Benefits:

a) Analytics on collaborative data is easier to deliver
b) Elastic computing capacity
c) Lower capital expense if leveraging a public cloud

Risks:

a) Increased data risk because considerable retail data may need to be loaded to the cloud
b) Up-front data load effort
c) Difficulty in switching providers once data have become entrenched with a particular provider

3.3.5 Electronic Commerce Service on a Public Cloud

Challenge: Multichannel retailing has become a core element of many retailers' operations, and an e-commerce channel is an essential way to reach consumers. Retailers need to give customers access over a public network (Internet) and provision computing capacity to meet peak demand to support e-commerce operations. If an e-commerce site goes down during peak demand, there can be a significant loss in revenue. Subsequently, complying with payment card industry (PCI) requirements is a concern for retailers. Other issues include the following:

- Capital investments
- Managing parallel channels
- Ongoing maintenance
- Security

Why Is a Cloud Solution Suitable? E-commerce is a natural fit for cloud-based solutions because of its ability to provide computing capacity on demand and to provision services for ubiquitous access.

Type of Cloud: Public or private

Type of Service: IaaS

Solution Brief: Retailers typically run their own Web sites and e-commerce channels, investing in the software, hardware, and human resources required to manage this channel. Moving this channel to a cloud platform improves the efficiency of the underlying hardware, storage, and networking infrastructure. Furthermore, if this channel is moved to either a dedicated third-party cloud host or to a public cloud platform, the retailer does not need to worry about provisioning for peak capacity and can instead pay for computing capacity based on usage patterns.

Benefits:

a) On-demand computing capacity

b) Centralized IT management of channel

Risks:

a) Availability of services

b) Accountability for loss of revenue or security breaches

3.3.6 Geo-Aware Tax Services on the Cloud

Challenge: Mobile retail has made the vision of "retail anywhere" a reality, and increasingly, through methods such as "pop-up" stores, retailers are able to go to where their customers are and sell their wares. E-commerce channels are also being increasingly targeted for taxation purposes, regardless of the difficulty associated with calculating taxes for any particular buyer. Tax calculation thus remains a specialized task with geo-specific or location-specific variations. It is simplest for most retailers to delegate this complex task to a service that can be accessed anywhere and can be used for calculating taxes in a geo-specific fashion.

Why Is a Cloud Solution Suitable? A tax service hosted on a cloud that offers universal access and can handle highly variable and seasonal demand is an ideal platform for such a service.

Type of Cloud: Public

Type of Service: SaaS and IaaS

Solution Brief: A tax service hosted on a public platform and offering standardized access to its consumers is the appropriate solution for supporting the needs of both e-commerce and mobile-based retailing. The tax service does not hold the transaction state, but instead computes the taxes and returns the value, acting much as a “task service” in SOA terms.

Benefits:

a) Standardized services that can serve a large number of retailers
b) Universal access
c) Computing capacity on demand

Risks:

The usual risks associated with cloud services: availability, reliability, and accountability

It is worth mentioning that there are several other services with similar characteristics to geo-aware tax services, with delivery routing services and weather-related services (as inputs to planogram planning) being two examples. These types of services are likely to gain a rapid foothold in the cloud environment for retail over time.

3.3.7 Private Cloud Test and Development Environments

Challenge: In some retail environments, significant investments in test and development environments remain underutilized periodically. The cost of setting up and tearing down these environments is high enough that retailers typically prefer to let those machines sit idle until their next period of use.

Why Is a Cloud Solution Suitable? Cloud technology permits the rapid automated provisioning and de-provisioning of services by moving entire images on- and offline quickly. This permits entire software configuration environments to be brought online and taken offline.

Type of Cloud: Private

Type of Service: IaaS

Solution Brief: Setting up an existing investment in hardware and software for use as a test and development environment in a cloud infrastructure or architecture permits the creation of virtual images of the required configuration for test and development. These images can then be moved online or offline reasonably quickly and at a low cost, allowing the underlying infrastructure to be used for provisioning new services.

Benefits:

Ability to reuse existing investments in hardware and software

Risks:

Cost of cutover from existing test and development environment to a setup based on a private cloud platform

3.3.8 Computationally Intensive Applications

Challenge: Many applications that help retailers make better buying and selling decisions, such as forecasting, planning, assortment, pricing, and replenishment applications, have relatively low ongoing operational computation needs (users interacting on a daily basis with the applications) and extremely high sporadic computation needs (typically weekly optimization or forecasting processes). In a traditional hardware environment, retailers are forced to procure hardware for the peak demand, even if most of this computational power is used for only several hours each week and is idle the rest of the time. As a practical matter, this has meant that some applications have remained beyond the reach of smaller retailers, who cannot afford the multimillion-dollar investment in hardware required.

Why Is a Cloud Solution Suitable? Cloud computing enables companies to match the computing power provisioned to actual needs at any given point in time, allowing companies to pay for modest computing power 95 percent of the time and to pay a small amount for much greater computing power on an as-needed basis.

Type of Cloud: Private or public

Type of Service: Primarily SaaS

Solution Brief: Running predictive retail applications such as forecasting, planning, assortment, pricing, and replenishment on the cloud allows retailers to run powerful simulation, forecasting, and optimization processes and pay only for the actual CPU hours of computing power consumed—a small fraction of the cost of procuring such hardware and having it be idle 95 percent of the time. During the 95 percent of the time that these same processes see a relatively light computational load (for example, a retailer or manufacturer entering initial plan numbers in a system, or defining new promotions to be run at a future date), the hardware cost is dramatically lower. Since most of the computationally intensive processes are run on a predictable batch basis, it is straightforward for consumers of cloud services to provision and de-provision computing power, and even to reserve peak computing power in advance, and thus take advantage of significant incentives (discounts) provided by cloud computing providers when customers can plan for and pre-reserve their computing needs.

Benefits:

a) Easy handling of predictable peaks in demand

b) Dramatically lower computing costs due to the provisioning of limited computing power 95 percent of the time and peak computing power 5 percent of the time

c) Availability of predictive applications to mid-sized and smaller retailers

Risks:

a) Connectivity and availability of cloud services

b) Data security for data living on the cloud (if using public cloud)

3.3.9 International and Smaller Retail Operations

Challenge: As retailers expand into new countries, or establish or acquire smaller operations, they often find themselves forced to use spreadsheets and manual efforts in many business processes for which the headquarters country has invested in software solutions. This reality is driven by the large up-front (and ongoing) expense, time, and effort required to replicate the hardware and software footprint of headquarters operations in each and every country or business. As a result, these retailers struggle to execute a consistent strategy and set of processes across their operations. Their international and smaller operations lack the right tools to run their businesses more effectively and efficiently.

Why Is a Cloud Solution Suitable? Cloud computing and related SaaS applications significantly reduce the expense, time, and effort required to roll out sophisticated systems to smaller operations.

Type of Cloud: Primarily public at present; can also be private

Type of Service: Primarily SaaS

Solution Brief: At an infrastructure level, a retailer's ability to use a public cloud service to deploy applications in multiple countries or across multiple businesses eliminates the need to invest in the real estate, hardware and software infrastructure, and operations expertise that would otherwise be necessary simply to run a data center. At an applications level, the ability to deploy various applications on an SaaS basis allows retailers to benefit from applications typically reserved for much larger installations (such as forecasting, planning, assortment, pricing, or replenishment) while paying only for the actual level of use of such applications in each geography or business. At a business level, by having these applications available across multiple geographies, retailers can ensure consistency in strategy and execution.

Benefits:

a) Rapid and cost-effective rollout of applications to international operations

b) Avoidance of investment in data center, hardware, software, and operations infrastructure for each operation

c) Provision of cost-effective access by smaller operations to enterprise-class solutions

Risks:

a) Connectivity and availability of cloud services across geographies

b) Data security for data living on the cloud (if using public cloud)

3.4 Additional Considerations

In concluding this chapter, a few additional comments are worth keeping in mind:

- The cost of each solution must be measured in terms of the impact to existing business and the cost of new cloud-based operations.
- There is a need to go against a preproduction environment during the cutover phase, something that has not been fully addressed by most cloud platforms today.
- One should keep the possibility of migrating services from a public cloud to a private cloud or IT infrastructure (and vice versa) in mind while making decisions regarding which services should be accessed in the cloud.

The following chapters address the cost of cloud computing as well as the chief concerns regarding integrating with the cloud and issues relating to reliability, availability, security, and accountability in a cloud-based solution.

4. THE COST OF CLOUD COMPUTING IN RETAIL

We have identified two broad categories of clouds, “public” and “private.” Public clouds are created by cloud providers to serve consumers from many different organizations on a public network. Private clouds offer the same computing platform, built on the retailer’s own infrastructure resources and serving the retail organization and any business partners or consumers the retailer offers access to. Both types of clouds strive to achieve more efficient use of physical IT resources by sharing or allocating them across more users.

Not all clouds fit all needs. The IT requirements for some companies may best be met by a combination of both public and private clouds. This will be particularly true as retailers look for cloud vendors that address specific areas such as replenishment, customer relationship management (CRM), sales taxes, or forecasting. Situations in which retailers need new capacity quickly or for a short time—for example, to support quality assurance (QA) or development activities—may be more suitable for public cloud utilization. Solutions that make use of both public and private clouds are called “hybrid clouds”; they are not a unique cloud structure but involve linking IT resources across both public and private clouds.

4.1 Public Cloud Costs

Public cloud computing migrates the retailer from a “build and own” model to a “rent and share” model. It eliminates up-front IT costs in exchange for ongoing service contracts. IT start-up costs move from the capital budget to the operating budget. This is very useful for start-ups and fast-growing retailers that cannot afford a large capital investment or do not yet know their final capacity requirements. It may make certain services affordable to smaller companies.

Public clouds can also be helpful in a rapidly evolving technology environment where delivery mechanisms are constantly changing and new interfaces must be supported. E-mail systems must talk to smart phones, and e-mail servers are integrated with voice mail. The public cloud insulates the retailer from technology changes and allows it to focus on core business processes. Public cloud providers can deliver online CRM and Web-store services with flexible branding options that create a custom appearance for each retailer while sharing the same technology stack. Public clouds can also deliver internal applications such as payroll, e-mail, and financial applications.

But as compelling as the cost savings are for cloud-based services, the cloud model threatens the old power structures of existing IT departments. Retailers have a large investment in existing IT infrastructure and personnel resources. These “sunk costs” are both a financial and political hurdle to using cloud-based services. The IT department may see public clouds as a threat to its own internal corporate influence. Critical operations may become vulnerable if cloud vendors cannot meet SLAs, and the retailer’s chief information officer (CIO) will be held accountable. If critical functions are moved into a cloud that uses proprietary software provided by the cloud vendor, the CIO may have difficulty transitioning to a new vendor that does not have the same features. Public cloud solutions may carry ongoing operating costs because they are not customized to specific retailer requirements.

The total cost of using a public cloud will depend on the range of services the provider offers. Infrastructure cloud costs are driven by CPU capacity, RAM, disk storage, and network bandwidth. Platform cloud costs include the operating system, database management system (DBMS), other system software, and data archiving. An application cloud supporting SaaS will often base costs on transaction levels, the number of accounts, or some other business measure. For example, sales tax or mapping services may have costs based on the number of tax jurisdictions or the geographic areas covered. When considering costs, growing retailers must insist on a clear explanation of how per-transaction costs will change as volume increases.

Since the whole premise of public cloud computing is to spread IT costs across many consumers, the cost of unique services that have fewer consumers will be higher than the cost of more broadly used services. An order fulfillment application that includes unique features for a particular retailer will have higher transaction costs than an e-mail application that is generic to all consumers. A typical public cloud business model is to offer a stripped-down version of a complex service for free (often supported by advertising) and a fee-based premium version with enhanced services, extended help desk support, faster response times, or less advertising. Retailers must consider whether a particular application is critical to their business and whether they need the higher-priced version.

International support for retailers that want to reach a worldwide market may be critical in the long term. IT support services may be combined with a fulfillment network that enables the retailer to reach customers beyond its normal market area. Companies such as Amazon combine not only the technical platform but also offer distribution centers in various locations around the world that can be used for fulfillment.

A simple comparison of in-house versus cloud implementations of an e-mail service was performed in 2008 by Merrill Lynch. The analysts estimated the cost to implement e-mail on a public cloud to be one-tenth the cost of hosting it on-site. They quoted other sources who estimated the public cloud cost to be one-fortieth of an on-site implementation. The largest factor reducing the cost was IT support personnel cost. Many of the personnel costs simply disappeared with the public cloud model. Table 2 shows the relative significance of various cost factors based on the type of implementation as they appear to the retailer. The challenge with this comparison is that it is like introducing a new store procedure that saves five supervision hours in the store. The store will not use fewer supervisors, though it might improve overall productivity (by allowing the supervisor to accomplish other tasks). This goes back to the “sunk cost” discussion above. (Platform costs exist in all circumstances but are concealed from the retailer in the service fees of the cloud vendor.)

Table 2: Cost of Public Cloud Computing: A Comparison

Cost Factor	Cloud-Based Implementation	On-Site Implementation
Requirement Specification	$$$	$$$
Climate-Controlled Environment	$	$$
Platform Selection (OS, DBMS, connectivity, etc.)	0	$$
Platform License Fees and Installation	0	$$
Application License Fees and Installation	0	$$
Application Configuration	0	$$
Enhanced Broadband Connection	$$	$
Backup and Recovery Plan	$	$$$
New User Training	$$	$
New User Licensing	$	0
Transaction Fees	$	0
Infrastructure Expansion	0	$$
Ongoing Software Updates	0	$

Table 2 notes:

- The requirements for either type of implementation must be defined by the retailer.
- A climate-controlled environment for the server, or at least a secure area for protection, is only necessary for an on-site implementation.
- Platform and application selection, licensing, and installation are only necessary for on-site implementations.
- Enhanced broadband connections may be necessary to maintain response times when services are moved onto the cloud.
- A backup and recovery plan is required in all cases but is included in the contractual arrangements with the cloud vendor. On-site implementations must install and test these procedures themselves.
- New user training may be somewhat more difficult for complex cloud-based applications that cannot be customized to each user.
- New user "seats" and transaction fees may be the pricing model for cloud-based services.
- Ongoing software updates are performed by the cloud vendor. This can be a double-edged sword if the retailer does not want or need a particular update and is forced to adopt (and pay for) it.

It is difficult to put a price tag on public cloud services because costs are constantly declining and vendors have different restrictions such as minimum usage or offer different services. Some of the cost factors to consider when evaluating vendors are as follows:

- Some cloud vendors have a minimum monthly charge. Conversely, some vendors offer free services up to a certain maximum, where charges begin.
- If a contract includes maximum usage limits, it should be clear what the fee rate will be if the business is so successful that the limits are exceeded.
- Windows instances on the cloud tend to be up to 20 percent more expensive than Linux or UNIX instances.
- SLAs seem to all be set at 99.95 percent as a standard. This translates to four inconvenient hours a year (which is in line with most in-house infrastructure).
- Retailers pay for flexibility. Costs come down if the company is able to commit to longer-term contracts or reserved usage levels.
- Rounding of usage times can affect total costs. Some vendors have minimum billing increments to which they round usage.
- Separation of computing and storage costs is important for part-time applications that need persistent data storage but can give up computing resources.
- Extended services may not be included in basic contracts. It may be necessary for a retailer to pay for a premium contract at least in the beginning until it becomes familiar with a particular cloud environment. The retailer should understand what it will cost to drop premium services later on.
- Storage costs may vary based on the number of files as well as raw data sizes.
- A retailer should be sure to have mechanisms in place to release resources when they are not needed. It might make sense to release testing or development resources overnight.

Moving to cloud-based services may also involve some preparation costs:

- Existing platforms may need to be standardized or consolidated in order to use the cloud-based platform. Existing applications that are running on different releases or versions of operating systems may need to be retested on the cloud platform.
- Before they are moved, it may be better to consolidate some services that are running on multiple instances.
- It will be necessary to test backup and recovery procedures by conducting recoveries and testing the recovered instances.

When evaluating the cost of a public cloud option, a retailer must determine the breakeven point, where transaction or seat costs begin to exceed the costs of an in-house solution. An established organization may not be able to achieve all the savings because technical skills must be retained for services that cannot be moved to the cloud.

The discussion below illustrates the kind of calculation a retailer must perform to determine its breakeven point. It describes how a cloud-based e-mail implementation might compare to an in-house solution. The retailer has an established IT department for other services. There is some variability in the cloud solution because it assumes the cloud vendor offers some type of quantity discounts. For the on-site implementation, it is assumed that the hardware is sufficient to support all the users. As the number of seats and length of time increase, the on-site cost per seat declines.

Assumptions: **Public Cloud**

Seats	Cost per Seat per Year
<26	**300**
26–50	**250**
51–100	**150**
>100	**100**

On-Site

Item	Cost in $
Hardware	**15,000**
License Fees	**90,000**
Communications	**8,000**
Annual Salaries	**100,000**
Physical Environment	**5,000**

One-Time	\$105,000 + 20,000	=	\$205,000
Ongoing	\$23,000 + 20,000 + 5,000	=	\$48,000

The average cost per seat on a public cloud for a retailer with 500 employees will be \$122.50 ((25*300+25*250+25*150+400*100)/500).

Assuming the hardware investment is good for five years, and the retailer's staff spends two months getting the system up and one day a week supporting it, then the private site would cost \$178 per seat ((205,000+5*48,000)/(500*5)).

The savings are not as great as Merrill Lynch outlined because we did not eliminate the salaries completely and the example is merely illustrative.

Table 3 provides costs from three cloud vendors as of August 2009. These numbers are only a starting point because they are constantly changing, and what each vendor includes in its various packages is not the same. There are other vendors providing cloud services as well, and the pricing in the table below is for illustrative purposes only.

Table 3: Public Cloud Costs from Select Vendors

Vendor	Resource	Resource Units	Cost per Unit in $
Google	Outgoing Bandwidth	Gigabytes	.12
	Incoming Bandwidth	Gigabytes	.10
	CPU Time	CPU Hours	.10
	Stored Data	Gigabytes/Month	.15
	Recipients E-mailed	Recipients	.0001
Microsoft	Data Trans In	Gigabytes	.10
	Data Trans Out	Gigabytes	.15
	Storage	Gigabytes/Month	.15
	Compute Time	Machine Hours	.12
	Storage Transacts	10K App Requests	.01
Amazon	Data Trans In	Gigabytes	.10
	Data Trans Out	Gigabytes	.14
	Storage	Gigabytes/Month	.15
	CPU Compute Time	Instance Hours	.125

4.2 Public Cloud: Indirect Benefits and Costs

The primary benefit of public clouds is the reduced dollar costs. These include hardware costs, infrastructure costs, and technical personnel costs. As with any technology, the benefits achieved by various organizations will vary widely depending on their starting points and their ability to absorb change. Public cloud implementations are much simpler for a start-up than for an established firm. Some of the additional benefits are outlined in Table 4.

Table 4: Indirect Benefits of Public Cloud Computing

Indirect Benefits	Description
Lower Up-Front Costs	Because the costs are fee based, the retailer can try new services without a large up-front commitment.
Faster Deployment	Faster deployment allows the retailer to maintain a leading edge on customer service.
Service Mashups	It may be possible to combine a variety of services from different cloud vendors or a single vendor in order to provide complete functionality.
Same Service from Multiple Providers	It may be possible to use different service providers in order to protect against "too big to fail" situations or to compare results for best outcome.
Accommodation of Fluctuating Demand	Retailers may be able to contract for lower rates during periods of low demand.
Reduced IT Footprint	Less equipment space is needed at the retail location, whether the store or the office. The challenge is that a robust network connection is necessary, especially at the store.
Fewer IT Personnel	With less hardware and infrastructure, the retailer requires less IT staff and management. But it may be difficult to achieve this benefit for existing organizations.

There are indirect costs associated with going to public cloud computing that are not merely dollars-and-cents costs. These costs are outlined in Table 5.

Table 5: Indirect Costs of Public Cloud Computing

Indirect Costs	Description
Reduced Utilization of Existing Resources	If the retailer already has “sunk costs” associated with an internal IT service, the cost of running the remaining applications could become prohibitive.
Forced Upgrades	The retailer may be forced to take upgrades and train personnel on new releases even though they do not provide significant new functionality.
Less Data Security	Cloud vendor personnel will have visibility of data in their storage facilities. Legal rulings have not yet made a precise determination of privacy rights for data stored outside a company’s premises.
Less Customization	Generalized solutions offered by public clouds may need to be adapted to a particular retailer’s needs by additional personnel training. (Heads-up for retailers: There are limits to configurability.)
Vendor Lock-In	Once a particular cloud vendor has all a retailer’s data, it becomes difficult for the retailer to change vendors. Even if SLAs are not met, the retailer may not have many alternatives.
Cutover Procedures	Methods are needed for testing and cutting over without impacting users. It may be necessary to support a parallel period while staff are trained and practice using the new system.

4.3 Private Cloud Costs

For organizations that already have an IT infrastructure in place, implementing cloud technology will allow them to speed delivery of IT resources to end users. Private clouds exist behind a corporation’s firewall. Instead of several weeks or months to process hardware requests; find floor space, cooling, and power; install the operating system, middleware, and application software; and lay the network and security, cloud technology presents IT services as a catalog of offerings from which users choose their resources. The cloud technology then provisions and allocates the resources automatically. User interfaces allow service users and IT administrators to fine-tune allocations and adjust configurations on a moment’s notice in order to meet actual requirements.

Private clouds are retailer-specific and reduce the physical hardware, environment, and software maintenance costs. They are supported by the retailer’s IT staff. Virtualization is a key component of cloud technology but it must be combined with administrative tools that speed allocation of resources. With private clouds, the time to provide new resources goes from days to minutes, reducing the hoarding of computing resources by user departments. Billing procedures can support metering of IT resources, which further encourages users to conserve them.

In a traditional environment, resources may be allocated to departments or applications. Because it is not always clear exactly what the final required capacity might be, resources are often over allocated and the unused resources may never be recaptured. Cloud technology introduces management software that monitors the virtualized environment to ensure efficient utilization. Key features of the management software are as follows:

- Automated Provisioning – Cloud computing supports user creation of new virtual servers and adjustment of physical resource allocation.

- Monitoring – Utilization of allocated resources is monitored and reported or even adjusted automatically.

- Capacity Planning – While cloud computing reduces the precision required for application-level planning, it increases the need for overall infrastructure planning.

Private cloud benefits vary widely depending on a company's starting point and final requirements. A typical implementation of a private cloud might be the virtualization of many corporate-based servers into a single hardware platform. This will reduce hardware costs, utility expenses, and floor space requirements. International Technology Group conducted an analysis of IBM virtualization costs. It determined that almost 75 percent of a large company's IT budget goes to infrastructure costs, and infrastructure savings can be as much as 50 percent on a private cloud. Countering these cost savings, the cloud requires additional software beyond virtualization software to manage the provisioning of resources.

Private clouds may be used for short-term resources associated with testing and piloting new releases or applications. Instead of expanding permanent resources, a retailer can use its cloud to provide temporary resources for development and QA instances of an application.

4.4 Private Cloud: Indirect Benefits and Costs

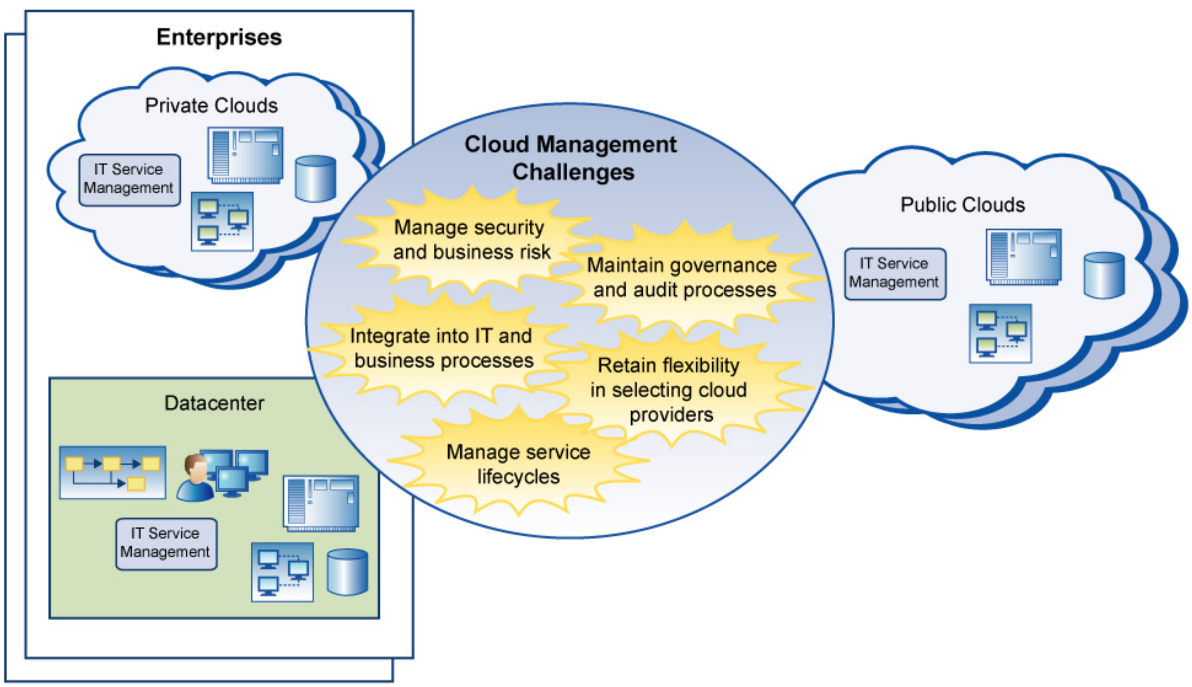
Because they are still proprietary and require retailer technical personnel for support, the primary benefit of private clouds is a reduction in hardware and infrastructure costs. Instead of each department having on-site servers, these allocations become merely partitions on a central server. Instead of maintaining a robust network between remote servers, less robust individual connections to a central server are possible. Consolidation and centralization of physical hardware reduces remote support calls.

For a retailer, centralization of in-store backroom processors can be a significant savings but may also require more robust network connections. An even more critical network decision would be to put point-of-sale (POS) support on the cloud.

5. INTEGRATING CLOUD SERVICES WITH RETAIL STORE AND ENTERPRISE IT

Services from a cloud frequently need to be integrated with services or applications running on a more traditional IT configuration within a retail enterprise to deliver valuable function to the business. However, a host of questions arise when attempting to integrate services from a public cloud, a private cloud, and a traditional retail enterprise.

Figure 7, reproduced from the publication "Interoperable Clouds" by the Distributed Management Task Force (DMTF), captures the range of issues that need to be addressed when dealing with integrating services from the cloud.



Source: DMTF, "Interoperable Clouds"

Figure 7: Management Complexity of Cloud Adoption

This section addresses some of the IT and business process integration challenges, while the next section focuses on security, governance, and accountability issues. The final section wraps up by discussing portability and interoperability issues.

Figure 8 is an example diagram depicting a high-level architecture in retail that includes services from a public cloud, a private cloud, the store, and the enterprise. Component position is prescriptive based on typical implementations.

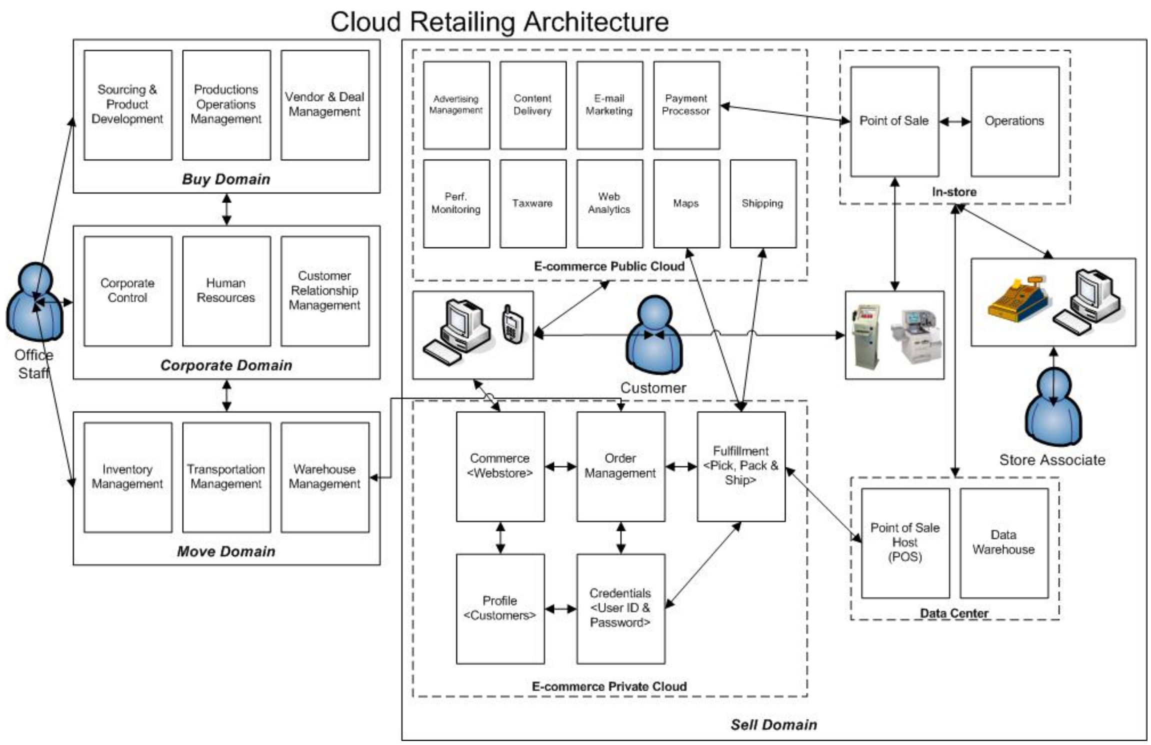


Figure 8: Integrating Services from the Cloud in Retail

The actual IT integration patterns can vary, as in any typical enterprise integration architecture; they can include but are not limited to the following:

- Process integration patterns (serial, parallel, etc.)
- Application integration patterns (message-oriented, broker-based, etc.)
- Data integration patterns (federation, population, etc.)

All levels of integration typically exist in retail environments, and they can all be developed using service-oriented principles to derive benefits at varying degrees.

When accessing services from the cloud, a few considerations must be addressed:

- Communication Latency – How long does it take for a service to be delivered, and is that appropriate for the application under consideration? For example, at the check-out line, delays beyond 20 seconds may be unacceptable.

- Data Transformation or Protocol Transformation – Does the data or service being delivered from the cloud needs further transformation before being integrated? This needs to be taken into account when choosing a service provider to keep integration patterns simpler.
- Idempotent or Reliable Nature of Service – While considering services from the cloud, it is helpful if the services are idempotent (give the same response to the same request) and, depending on the nature of the integration, offer some form of guaranteed or once-only delivery.
- Security and Integrity of Data Transmission – Data integrity and security must be ensured while accessing services from the cloud.

Furthermore, choreography and orchestration remain major considerations for accessing services on a cloud, whether public or private. Choreography is dependent on each service's role, given that there can be multiple clients and services. Consideration should be paid to service interaction because choreography is premised upon the sum of the parts, or it may be process-oriented, wherein control resides with a single component (i.e., service). For optimization purposes, a long-running transaction may need to be broken into smaller pieces. Other options include moving the entire transaction to the cloud.

Many of the typical considerations are thus those that one would account for in any enterprise architecture, and the more service-oriented the architecture is, the more flexible the integration would be.

5.1 Moving Services to the Cloud: Service Categories

The ARTS SOA Best Practices Technical Report classified services into categories, and that approach is helpful in determining what should move to the cloud for ease of integration with the enterprise:

- Utility – Stateless auditing, logging, and other utilities typically do not contain business logic and are excellent migration candidates. These utilities are among the easiest services to migrate and integrate with the rest of IT.
- Task – Self-sufficient and stateless, a task involves a complete transaction on the cloud by passing data in, applying business rules, and receiving data back.
- Entity – Data can be managed in the cloud for entities such as customer, vendor, and so on. Since these entities can often involve sensitive data, they may not be good preliminary candidates for moving to the cloud.
- Process – A process service usually carries state as it uses other services. Determining which process services to move to the cloud should include mapping existing relationships.

Ultimately, a retailer needs to create and complete a checklist with a multidimensional matrix before moving services to the cloud. Some entities are not good candidates because highly confidential data could be involved. However, nutrition information for items, for example, is an excellent entity candidate.

5.2 Examples of Integration with Services from the Cloud

Table 6 contains some examples of integrating services on private and public clouds.

Table 6: Examples of Integration with Cloud Services

Example Description	Hybrid	Private	Public
Using a hybrid cloud service to transform proprietary POS logs to the ARTS POSlog system as well as communicating with external vendors	●		
Mashing up private and public services to provide fresh-item management	●		
Creating an XML hub for multiple POS systems with different transaction logs		●	
Credit and debit payment system authorization transactions			●
E-commerce, including multichannel (e.g., kiosk), transactions that typically rely heavily on a wide variety of cloud services (A cloud-based shopping cart is an example of a process-classified service.)			●
Providing customers digital receipts by interfacing POS with public services			●

Migrating services to or leveraging services from the cloud means not only integration challenges to consider, but also security, reliability, availability, and accountability issues to overcome. The next chapter addresses some of these top concerns with cloud computing.

6. SECURITY, RELIABILITY, AVAILABILITY, ACCOUNTABILITY: OVERCOMING THE TOP CONCERNS WITH CLOUD COMPUTING IN RETAIL

Cloud computing typically raises concerns regarding security, accountability, availability, and reliability. The promise of scalability is one of the compelling features of cloud computing and needs to be verifiable to prevent a letdown in the midst of peak demand. It is also important to understand the role of “risk acceptance” in the decision-making process for the adoption of any technology in retail, including cloud computing. When adopting any new technology, trade-offs are made with regard to security, availability, and reliability—depending on the critical nature of the operation being supported, the quality of service required, the associated performance needs, and other considerations. Adoption of cloud-based services needs to be seen in the same light, and appropriate actions may need to be taken, depending on the part of the retail operation that is being served by the cloud service.

6.1 Cloud Computing Security

Security aspects of private clouds are similar to those of enterprise data centers. Public clouds, in contrast, are associated with a particular set of security issues. A highly distributed service-based model renders many existing enterprise security architectures obsolete. Security architects from all major cloud computing vendors are in the process of reengineering security models that can be implemented in a distributed cloud infrastructure. The traditional “defense-in-depth” approach to security must be extended beyond enterprise data centers to distributed and federated environments. New solutions should be flexible enough to work with different cloud implementations. Even though some risk can be transferred to the cloud, all the major issues related to accountability and responsibility for protection of sensitive data still rest with the original owners of the data. Retailers that plan to use cloud services should clearly understand and assess these potential vulnerabilities. In many cases it might be a good idea to request a security risk assessment from an independent third-party firm specializing in these matters.

Retailers should avoid vendors that do not provide clear and detailed information on their security measures, practices, and implementation. It is important to recognize that there is no such thing as absolute security. It is always a balance between acceptable levels of risk and the cost of implementing and managing different security measures.

Cloud-specific security concerns can be roughly divided into two major groups: issues associated with managing data outside a retailer’s private networks and issues related to the public nature of services exposed over the Internet.

6.1.1 Private Data in the Public Cloud

Different types of data within the retail enterprise have different levels of sensitivity. Some data, such as tax calculation rules, item information, and electronic coupons, are not very sensitive, whereas the personal records of customers and employees and financial records are examples of data that have strong privacy requirements. So, it is only natural that many retailers feel more comfortable putting the types of data that have fewer privacy concerns into the cloud.

When private data are uploaded to a public cloud, they are potentially exposed to maintenance and support personnel who have not gone through a retailer's internal data security management programs. The retailer cannot perform background checks or provide any oversight of the staff with access rights to its sensitive information. It is also important to find out if there are any third parties that the cloud service provider contracts with to perform maintenance tasks and what kind of privileged access they might have to a retailer's data. Physical protection of the data is essential. It is also a good indicator of how seriously cloud platform and/or infrastructure providers treat the issues of data access security. Therefore, it is a best practice to inquire about a vendor's physical safety measures, such as, "Do data center facilities have any boundary protection such as perimeter-control berms, fences, and so on?" It is important to make sure that physical access is strictly controlled through multiple checkpoints and that appropriate video surveillance and intrusion detection systems are used to safeguard access to the data center. It is often found that vendors' security programs are more stringent and robust than the programs used internally by many retailers. Additionally, retailers may have a stronger legal recourse to recover damages caused by a security breach in the cloud than they would have had if the infringement had happened in their own data center.

Figure 9 illustrates the importance of contracts and SLAs related to private data on public clouds.

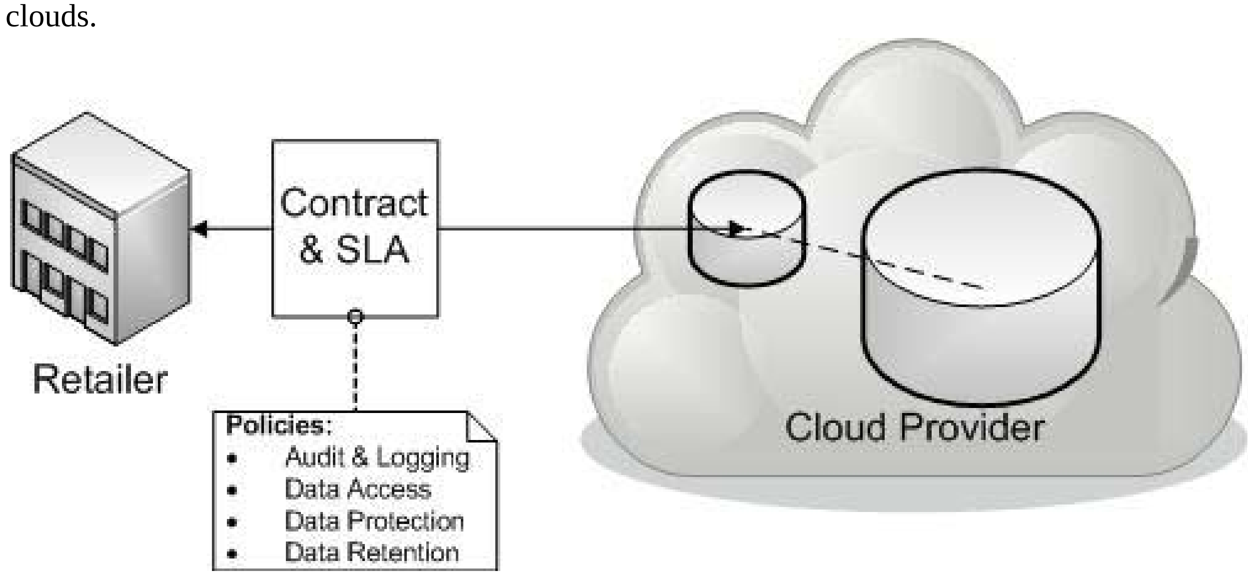


Figure 9: Private Data in Public Clouds

Data that reside within the cloud must be treated in compliance with industry and government regulatory requirements. For example, many retailers have invested a great deal of time and money to achieve PCI compliance. To remain PCI compliant when sensitive cardholder information, such as credit card numbers, is uploaded to the cloud, a retailer must have contractual guarantees from cloud service providers that their systems are PCI compliant and that they will bear full responsibility for the data in the cloud (see Section 6.4, Cloud Computing and PCI, for more information). There are often other regulatory requirements, such as the physical location of the data in the cloud or the maximum time period that the data can be kept, to consider as well. Retailers should ask where the data are physically kept and inquire about the details of data protection laws in the relevant jurisdictions. Retailers typically have only limited control over the physical aspects of the data in the cloud, and they should seek contractual commitments from cloud service and infrastructure providers that these regulatory and compliance requirements will be satisfied.

A typical characteristic of cloud computing is multi-tenancy, in which a single software system running in the cloud serves multiple client organizations (tenants). This feature introduces additional security concerns and requires strict data segregation. If a cloud service provider uses multitenant architecture, retailers should request information about data segregation procedures and guarantees.

In addition to contractual guarantees from cloud service providers, retailers can use techniques such as encryption and tokenization to alleviate some of the data security concerns mentioned above. These techniques can be an effective last line of defense and represent an important part of the multilevel data protection architecture. If such techniques are enabled through services within the cloud, service providers should also give guarantees that encryption and tokenization schemas are robust and properly tested.

An important issue associated with retailers' data being stored in the cloud is disaster recovery. Retailers regularly perform data maintenance procedures such as backups, archiving, and so on to ensure that important business information can be recovered in case of a system failure. The cloud service provider should be able to describe clearly what will happen to the retailer's data in the event of a disaster. The retailer should ask specific questions about guaranteed and complete data recovery. Typically, robust data protection involves data replication across multiple physical devices and even multiple sites. Retailers should inquire about recovery time frames and potential impact on service availability.

When a retailer's data resides in a public cloud, monitoring and maintenance tools offered by the cloud provider's platform become extremely important. It is essential that the cloud service provider can commit to an appropriate level of support for investigations that a retailer might need to perform to analyze suspicious activity.

In addition to the technical issues of data security outlined above, it is also essential to consider business aspects such as the viability of the cloud service provider as a company. If the service provider fails or is acquired by another company, what will happen to the retailer’s data? A second business aspect to consider is data ownership and legal protection of the retailer’s data in the cloud. The data represents the retailer’s property, but it may have less legal protection against search and seizure when it is uploaded to the cloud. The government or an attorney performing discovery might be able to gain access to the data more easily than if it were stored on the premises inside the retail enterprise.

The subject of data ownership is primarily comprised of two differing views: an “internal administrative” view and an “external legal” view. The internal view, often dominated by the retailer’s information security department, is focused on data governance within an enterprise. Its focus is policy and procedural topics such as data classification, accountability, access control, data validation and data quality, data life cycle events, and so on. It often follows information security standards such as ISO 27001. The external view, in which governance moves from internal information security policy to also include the retailer’s legal department, is concerned with the rights and obligations of different legal parties, including the author or originator, editor, distributor, and consumers. When contracting with cloud service providers, the definition of data ownership should be explicitly spelled out and include the roles and responsibilities of all parties. As such, data ownership is more about responsibilities than rights. For example, in the ISO/IEC 27001 standard, ownership is explained as follows: “The term ‘owner’ identifies an individual or entity that has approved management responsibility for controlling the production, development, maintenance, use and security of the assets. The term ‘owner’ does not mean that the person actually has any property rights to the asset.”

There have been no significant challenges, such as lawsuits, involving cloud-based data yet, possibly because there have been few, if any, major breaches of service or security (at least publicly reported). In this context, forensic capabilities are another important area to inquire about when researching cloud-based service providers. Forensic tools offer the ability to monitor in real time as well as assess after the fact what occurred in the event of a service compromise or attack. There are a growing number of tools and packages available for private networks to determine the breadth and extent of an incident, but for public offerings the tools and methods needed to collect and process cloud-based forensics are only beginning to be defined. The need for proper risk assessment in areas such as data integrity, recovery, and privacy, and the evaluation of legal issues in areas such as e-discovery, regulatory compliance, and auditing, are crucial when transitioning to cloud solutions.

When storing data in the cloud, retailers often do not know exactly where the data are physically hosted, or even what country they are stored in. A best practice is to contractually oblige service providers to store and process data in specific jurisdictions and to obey local privacy requirements on behalf of the retailer.

6.1.2 Cloud Service Security and Accountability

One of the major selling points of the cloud computing approach is that employees and partners can access critical applications from virtually anywhere. Unfortunately, this also exposes services to potential security threats and increases the risk of unauthorized access.

Access to sensitive data outside the retailers' enterprise brings with it an inherent level of risk, as cloud services can bypass the physical zoning controls retailers use with their existing systems. Retailers should obtain as much information as they can about the people who manage their data by asking cloud service providers to supply specific information on the hiring and oversight of privileged administrators, and the controls over their access.

The first aspect of security is authentication. To protect the retailer's assets in the cloud against unauthorized access, strong authentication solutions are essential. Unfortunately, un-guessable passwords also tend to be un-memorizable. As a result, they are extremely inconvenient and difficult to use. Multiple examples of security breaches on the Internet demonstrate the importance of strong authentication. One method to address this issue is implementation of two-factor authentication. This technique requires two different pieces of information to be used in conjunction to authenticate the user. A typical approach would be to use "something the user knows" and "something the user has"—for example, a PIN and a physical credit card, or a password and a digital certificate. A more user-friendly mechanism would be to utilize single sign-on technologies from the federated identity area such as OAuth, OpenID, SAML, InfoCard, and so on. Unfortunately, there is no agreement between cloud vendors on a standard way of handling universal IDs. At this point, the problems with federated identity on the Internet have not been solved in the standards. In the future, cloud customers will expect that services exposed on the cloud comprise an interoperable and secure environment.

Providing maximum flexibility for connection options reduces control over network security. In order to secure a connection over an unknown network, retailers that use cloud computing should apply some level of transport security that protects data. The path between the client and the service deployed in the cloud provides an opportunity for an attacker to steal information via passive sniffing or a more active man-in-the-middle or replay type of attack. Because the underlying public network cannot be fully trusted, using virtual private networks (VPNs) with secure sockets layer (SSL) and IPSec for transport security can offer access to all appropriate data center resources in a secure way. A VPN approach ensures that end users are compliant with the vendor's and the retailer's security policies through their implementation of endpoint security. As a practical matter, SSL VPNs are often cheaper to implement and operate than "conventional" IPSec-based VPNs. As SSL is built into every Web browser, SSL VPNs do not require special client software installed on every device communicating with the cloud-based service.

As cloud computing matures and is able to address more and more of the security concerns mentioned above, it should become a common platform for deploying an increasing number of services. It is conceivably easier to maintain a highly secure infrastructure environment with a relatively small number of cloud service providers than an on-premises environment of an organization. Many companies already use cloud solutions today, often without realizing it. For example, services such as BlackBerry RIM involve messages being sent to a third party, scanned, and then forwarded back into the organization's localized e-mail stream. This demonstrates that a well-run cloud solution can overcome the perception of insecurity and loss of control, and that it can be securely implemented.

6.2 Availability

Availability of a service means that it is up and running and responds properly in a timely fashion, as defined by an SLA.

Monitoring service capability becomes an important part of ensuring high availability. This is especially true in a volatile environment such as cloud computing, where service instances can be provisioned and de-provisioned on a fairly frequent basis based on the level of activity. Detecting that a service is not responding properly is only the first part of the solution. The second part is to take appropriate measures to correct the problem, such as retrying the request against another service instance or sending a request to the provisioning system to launch a new instance of the service.

Even though the service might be available in the cloud, it might not be available to a client application running on-premises in a store. The mere fact that cloud services are remote creates the possibility of them being unreachable. Many potential difficulties exist along the path to a cloud service, such as internal issues with a retailer's network or problems with ISP connectivity, which can negatively impact service availability.

These simple facts have two practical implications. First, it does not make sense to incur extra costs for very high levels of service availability in the cloud if other elements of the communication infrastructure are not capable of delivering this high availability to an end client running on-premises. If the retailer does not have a highly available Internet connection, it does not matter if the service on the cloud has five nines availability (99.999 percent, only about six seconds of downtime per week). Second, it would either restrict retailers to deploying cloud services that are not mission critical, where some downtime would be acceptable, or they would have to design client applications that would handle an offline condition gracefully with potential failover to a service running locally.

The availability of a service depends security and network resiliency, as discussed above. For example, distributed denial of service (DDoS) attacks is specifically designed to render services unavailable. Also, if the service data are corrupted, the service will not be able to respond properly, thereby also making it unavailable. Thus, an unsecure service cannot be considered highly available.

6.3 Reliability

Reliability is the combination of a service's accessibility (the service is available when desired), continuity (the customer has uninterrupted service over the desired duration), and performance (the service is able to meet customers' expectations). Security and availability, discussed above, tie directly to these key aspects of cloud service reliability. Availability is often measured as the percentage of time that a service is up and running and responds properly, according to the SLA, and as such is the major characteristic of service reliability. Of course, it is difficult to consider a service reliable if it is vulnerable to security attacks.

Specifically for cloud computing, reliability is a measure of how well a service can cope with the challenges of a distributed environment. For example, if a service becomes unavailable even within the limitations of the SLA, would it cause any undesirable side effects?

Highly distributed systems must be designed differently. It is impossible to eliminate all undesired phenomena, but services can be designed to successfully cope with them. The ability of the system to deal with such abnormalities makes it more reliable. For example, data to cloud services should be sent with an "at least once" delivery guarantee rather than the "exactly once" delivery expectations typical in monolithic systems. It means that cloud services must be designed with idempotency in mind and should not rely on a particular order of messages. The use of "at least once" delivery guarantees combined with idempotency is logically equivalent to "exactly once" delivery with highly improved reliability. This means that a system designed in such manner should behave reliably even if connectivity is temporary lost.

It is worth highlighting additional cloud computing concerns, such as forensics and audit logging. It is conceivable that cloud computing can make forensic investigations easier, but this is true only if systems are deployed and used properly. Being able to provision new servers quickly is easy when responding to an incident, but the merchant can just as easily decommission compromised servers improperly and lose valuable information.

6.4 Cloud Computing and PCI

The cloud computing models all differ with regard to Payment Card Industry (PCI) compliance. Both private and public clouds may handle the processing of cardholder data differently based on how transaction processing is performed within the cloud, or where cardholder data are stored or pass through systems in the cloud.

6.4.1 Private Clouds

Computing in private clouds places all responsibility on the retailer to maintain systems and services in accordance with PCI requirements. Segmentation of data and systems can be utilized to help reduce the scope of the cardholder data environment but may limit the flexibility that cloud computing can provide. As both the PCI Data Security Standard (DSS) and Payment Application DSS set a minimum bar for security best practices, it is a good basis to start with for securing the entire enterprise.

Key areas that a PCI auditor has in mind when evaluating retailer compliance should include proof of the following[1]:

- The "cardholder data environment" can be segregated from other systems.
- The retailer can define and control access to "authorized" systems and users.
- The "ownership" of cardholder data can be tracked through the cloud.
- The retailer has change controls for the hypervisor and virtual infrastructure (VI).
- The network isolation of hypervisors from guest OSs is via an admin segment.
- The retailer can detect and limit the risk due to misconfigured virtual machines (VMs) or systems.
- The intrusion detection system (IDS) has the ability to monitor traffic between VMs and within the cloud.
- The file integrity monitoring tools can detect changes in active VMs.
- The patch management includes patching the hypervisor and the VMs.
- The retailer can detect changes in VM consoles, super admin IDs, and parent VMs.
- The retailer can monitor and block "rogue" VM IDs, including super admin IDs.
- The VI is auditable as part of the change management process.
- The retailer has segregated VM and/or cloud administrative duties.
- The retailer can identify and quarantine noncompliant VMs or systems.
- The retailer does not share network interface points among VMs or systems.
- The host OS (especially Windows, Solaris, or Linux) is properly hardened.

In general, assessors want proof that software and administrative controls have been implemented to work in a cloud environment. By showing segmentation, security monitoring, and tracking of all access to the "cardholder data environment," compliance can be achieved.

6.4.2 Public Cloud Services

Retailers must ensure that their service providers are PCI DSS compliant. Cloud computing using publically available services falls specifically under PCI requirement 12.8: "Sharing card holder data with service providers." The PCI DSS itemizes four sub-

[1] Provided by David Taylor of PCI Knowledge Base

requirements that must be followed and verified for compliance. For retailers performing their own self-assessments, there is a bit of a gap, as the simplified self-assessment questionnaires do not call out all the specific PCI requirements that are included in a formal audit as performed by a Qualified Security Assessor (QSA). It is the responsibility of the retailer to ensure that it is compliant with all the PCI DSS requirements.

Using cloud-based services can help reduce the scope of a retailer's PCI compliance infrastructure. Instead of creating in-house payment capabilities and going through the PCI compliance efforts, these services can be contracted from the cloud by service providers that are already compliant. Subscription services providers in the cloud that are already taking credit card payments must be PCI compliant.

There are generally two ways to be compliant, based on how a retailer handles credit card and billing information when using public cloud services:

1. The merchant accepts, handles, stores, and processes all cardholder information. With a decision to host the forms and services that collect cardholder data, store it, and manage recurring transactions, the merchant will need to implement all the physical security, network security, and application security required by the PCI standards. If services are hosted by a cloud provider that cannot or will not meet these requirements, then the merchant will not be able to become compliant.

2. A PCI-compliant service provider is used to collect, store, and process all subscriber cardholder data. When a service provider is used, the merchant's commerce site does not need to collect or process any cardholder data. While the site is still required to become PCI compliant, the scope of compliance can be significantly reduced and may only involve filling out a PCI self-assessment form. In this form the service provider is identified as handling all cardholder data. For these types of agreements, the retailer should ensure that the service provider complies with the following:
    - The service provider has service-provider-level PCI compliance.
    - The service provider's application or portal never allows access to subscribers' cardholder information.

6.5 Select Cloud Computing Security Developments

Cloud services are growing to meet enterprise needs for security, and new cloud-focused security certifications are on the horizon. These new cloud security certifications are expected to go beyond existing Statement of Auditing Standards 70 (SAS 70) and International Organization for Standardization (ISO) 27001 standards that service providers identify today as their security certifications. These existing standards are insufficient for assuring potential cloud customers that the cloud service provider has deployed proper security or that their data are adequately protected.

SAS 70 certification is typically viewed as a foundational certification. It is fairly open-ended in that the service provider defines the controls and an auditor then makes a judgment call as to whether these controls are sufficient. Each service provider's controls are unique, with little comparison of the comprehensiveness between providers. SAS 70 certification is clearly inadequate in certifying complex cloud-based systems. The PCI DSS is a better standard, as is stipulates a series of controls and how they should be implemented and includes both logging and monitoring.

ISO 27001 is a much broader specification than SAS 70 and as such is a better fit for cloud services at this time. ISO 27001 specifies how an organization should handle its information security management, security controls, and risk assessment.

Currently, few cloud providers state that they have either a SAS 70 or an ISO 27001 certification, and among those that do, SAS 70 is the most common. More cloud customers are demanding to know how their sensitive data are being protected. Cloud service providers need to implement these and future security certification standards to strengthen the viability of their service offerings.

The Cloud Security Alliance is one group working with other key players in cloud security and auditing to determine which organizations should provide new cloud-specific certifications, as well as what such certifications should include. A statement of direction is expected in the first quarter of 2010.

With regard to PCI, a special interest group (SIG) has been formed within the PCI counsel to provide guidance on virtualization. Current discussion concerns the host level but is being extended to the larger network-services level. This guidance is also expected to extend to clouds.

The potential for fraud is a major inhibitor to enterprises and their users in embracing cloud services. Retailers need to expand their strong authentication and fraud-detection capabilities so as to protect against unauthorized access, phishing, malware, and even intellectual property theft for both public and private clouds. For security experts, a recently released RSA Security Brief offers specific advice on how to best implement identity and data protection in the cloud. Additionally, RSA's solutions and services center has detailed information on data center monitoring and multitenancy, data encryption and tokenization, federated identity management, strong risk-based authentication, fraud prevention and malware detection, cloud event management and audit, data loss prevention, and regulatory compliance.

In addition to security, vendor lock-in and portability of services are key concerns regarding cloud adoption. The next section begins to look into the subject of an open cloud.

7. AVOIDING VENDOR LOCK-IN: SUPPORTING AN OPEN CLOUD

As cloud services become increasingly available from a number of software vendors, the potential for development of proprietary methods of access, and hence vendor lock-in, increases. As an open industry standards organization, ARTS promotes an approach to cloud computing that will ensure that service implementations are based on open standards so that retailers can avoid being locked in with proprietary vendor solutions. Retailers and providers using cloud services that conform to a standards-based solution will gain the benefits of interoperability, portability, and manageability.

Leveraging existing ARTS SOA Blueprint and Best Practices recommendations as well as publicly agreed-to principles of cloud computing presented in works such as the Open Cloud Initiative, this section will outline recommendations specifically for retail cloud solutions.

It is important to note that developing an open-standards-based recommendation is not a “start from scratch” piece of work. Hundreds of retailers worldwide have already adopted the principles of service-oriented architecture (SOA). SOA’s architectural approach provides the flexibility and agility required by today’s global business environment. In developing a position on open standards in cloud solutions, it is important to note that cloud environments will derive significant benefits from leveraging existing SOA standards. This is our starting place in developing an ARTS position for cloud computing.

7.1 Key Drivers for an Open Cloud

As retailers consider the use of cloud computing to offload specific services to a private, public, or hybrid cloud, it is important to understand the key drivers that will ensure an “open cloud.” When cloud computing solutions support interoperability, manageability, and portability, retailers, as cloud consumers, will benefit from a truly open cloud solution that is flexible and cost-effective.

In this section, we will discuss what it means for a cloud solution to be interoperable, manageable, and portable.

7.1.1 Interoperability

Interoperability refers to the ability of services on the cloud to communicate with the service consumer and with other cloud services (Figures 10–12). In this section, we identify how standards may be used to enable interoperability between various combinations of service consumers and services on the cloud.

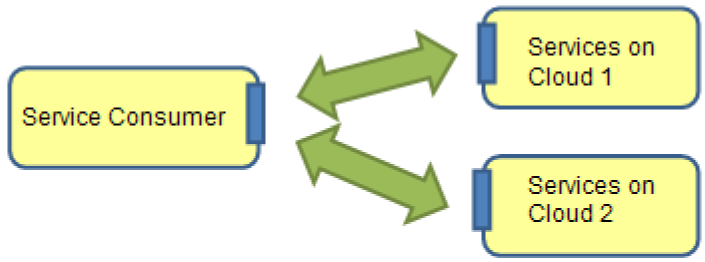


Figure 10: Consuming Services from Multiple Cloud Platforms

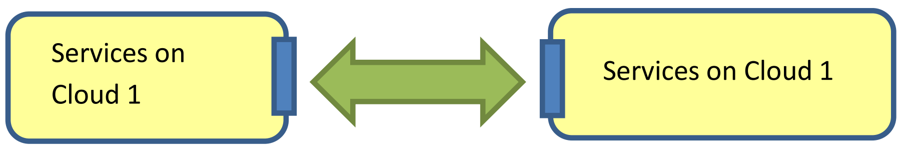


Figure 11: Interoperability between Services on the Same Cloud Platform

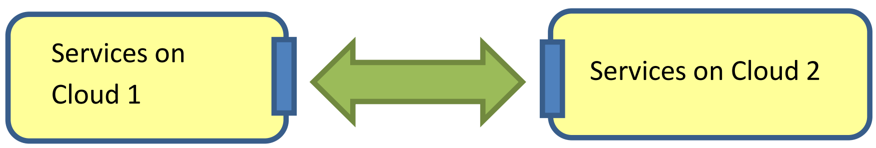


Figure 12: Interoperability between Services on Different Cloud Platforms

Services provided by various cloud vendors utilizing open standards will ensure interoperability and will gain the benefits that such an approach provides for the interactions between them. Open standards that should be considered for interoperability include those for transport, message format and protocols, identity or location, quality of service, and service information access.

Organizations will gain maximum flexibility for creating new solutions only when data and applications interoperate with each other regardless of where they reside (public clouds, private clouds that reside within an organization's firewall, traditional IT environments, or some combination of these). Cloud providers need to support key interoperability standards so that organizations can combine any cloud provider's capabilities into their solutions.

7.1.2 Manageability

Manageability relates to integrating and managing applications or services in a cloud environment (Figure 13). Management is an important aspect of SOA, and an approach based on open standards enables consistent manageability. Service consumers using cloud capabilities that conform to open standards gain the benefits of uniformity that such an approach provides for managing the interactions between them. It is important to note, however, that a private cloud (i.e., "cloud inside the enterprise") is a space where retailers have to worry even more about manageability.

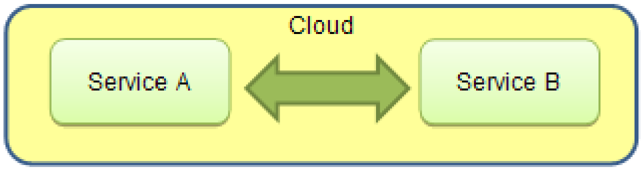


Figure 13: Integrating Services on the Same Cloud Platform

Open-standards-based manageability provides the capability to configure, monitor, manage, and maintain the cloud environment. To ensure manageability, open standards should be used for the following:

- Data Repositories – Information describing cloud capabilities is maintained by providing a registry, which serves as a directory for identifying and locating services, and a repository maintains the description of the requestor and provider service requirements and service capabilities.

- Service Definition – Standards should be used to describe service interfaces (e.g., WSDL) as well as data format (e.g., XML schemas). XML provides a platform-neutral mechanism for transferring data between services.

- Information Model – Information models should be used for expressing document content (XML) as well as expressing document semantics. Information models also provide definitions for management information for systems, networks, applications, and services.

- Service Characteristics – Metadata describe characteristics of a service and requirements associated with the use of a service. Other information may be used to characterize service definition with additional semantic annotation. Additionally, SLAs are needed to define the contract between the parties (consumer and provider) to establish common understanding about services, priorities, responsibilities, guarantees, and warranties.

- Service Federation – The repository standards discussed above could be used to share the metadata definitions across multiple cloud providers. The information would be used by the "federated" cloud to provide namespace mapping, supply the appropriate service interaction across multiple implementations, and determine the required subset of services applicable throughout the enterprise.

- Metadata Access – A standard format should be used for packaging and describing VMs and applications for deployment across heterogeneous virtualization platforms—for example, DMTF's Open Virtualization Format (OVF).

- Content Management – Content management interoperability services (CMIS) provide interoperability with enterprise content management (ECM) systems by leveraging open standards, including SOAP (simple object access protocol), REST (representational state transfer), and Atom.

7.1.3 Portability

Portability (Figure 14) is concerned with the ease of moving components or systems between cloud environments (hardware and/or software environments). Portability relates to the programming model aspects used within the cloud, including programming interfaces, information content, and core requisite facilities (e.g., provisioning, storage, messaging, datasets).

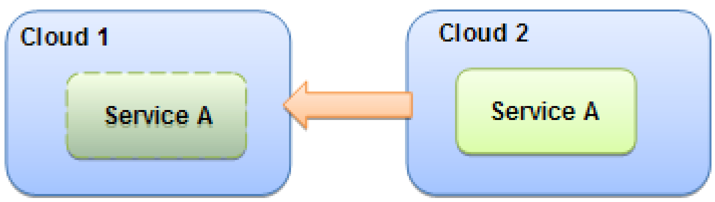


Figure 14: Moving Service from One Cloud Platform to Another

It is important to note that without standards, the ability to bring systems back in-house or choose another cloud provider will be limited by proprietary interfaces. Once an organization builds or ports a system to use a cloud provider's proprietary offerings, bringing that system back in-house will be difficult and expensive.

7.2 Organizations Addressing Cloud Standards

There are quite a few organizations currently addressing open standards for cloud computing. The following is a list of those organizations, along with an explanation of the cloud standards aspects addressed:

- Open Cloud Manifesto (http://www.opencloudmanifesto.org) – The Open Cloud Manifesto is a statement of principles for openness in cloud computing. The goal is to raise awareness and educate customers on vendor lock-in and proprietary cloud technologies. The Open Cloud Manifesto is a catalyst for standards organization coordination.

- Cloud Computing Use Case Discussion Group (http://groups.google.com/group/cloud-computing-use-cases) – This is an initiative that originated with the Open Cloud Manifesto community but has expanded far beyond its original supporters. It is an online experiment in group authorship utilizing an open source process. The objective is to produce a white paper that describes the high-level use cases and requirements for an open cloud environment based on real customer requirements.
- Simple Cloud API (http://www.simplecloudapi.org) – Simple Cloud API is a new open source initiative that enables developers to build truly portable cloud applications. Developers can have their software applications interface to a variety of cloud environments without making time-consuming and expensive changes to their source code, giving choice and flexibility in cloud providers.
- DMTF Cloud Incubator (http://www.dmtf.org/about/cloud-incubator) – The DMTF's Open Cloud Standards Incubator will develop a suite of DMTF informational specifications that deliver architectural semantics to unify the interoperable management of enterprise computing and cloud computing. This may include extensions to existing DMTF specifications, including the Common Information Model (CIM) and Open Virtualization Format (OVF); member submissions; and investigation of opportunities for collaboration with other industry standards bodies.
- SNIA Cloud Storage (http://www.snia.org/cloud) – The Cloud Storage Technical Work Group was created for the purpose of developing Storage Networking Industry Association (SNIA) architecture related to system implementations of cloud storage technology.
- Open Cloud Consortium (http://www.opencloudconsortium.org) – The Open Cloud Consortium (OCC) is a member-driven organization that supports the development of standards for cloud computing and frameworks for interoperating between clouds, develops benchmarks for cloud computing, supports reference implementations for cloud computing, manages a test bed for cloud computing called the Open Cloud Testbed, and sponsors workshops and other events related to cloud computing.
- Open Cloud Computing Interface (http://www.occi-wg.org) – The Open Cloud Computing Interface (OCCI) is a multivendor initiative to deliver a royalty-free, open standard API. OCCI will focus on the creation of an API for interfacing IaaS cloud computing facilities that is sufficiently complete to allow for the creation of interoperable implementations.
- Cloud Security Alliance (http://www.cloudsecurityalliance.org) – The Cloud Security Alliance's (CSA's) mission is to promote the use of best practices for providing security assurance within cloud computing, and to provide education on the uses of cloud computing to help secure all other forms of computing.
- The Open Group Cloud Computing (http://www.opengroup.org/cloudcomputing/) – The Open Group has formed a cloud computing workgroup to create a common understanding among buyers and suppliers of how enterprises of all sizes and

scales of operation can include cloud computing technology in a safe and secure way in their architectures to realize its significant cost, scalability, and agility benefits.

8. REFERENCED DOCUMENTS

Open Cloud Manifesto	http://www.opencloudmanifesto.org/
Open Cloud Initiative	http://www.opencloudinitiative.com
Open Cloud Principles	http://www.opencloudinitiative.com/ocp
VMware Write-Up on Internal Cloud	http://www.vmware.com/files/pdf/cloud/eight-key-ingredients-building-internal-cloud.pdf
VMware White Papers	http://www.vmware.com/solutions/cloud-computing/resources.html
Five Companies Shaping Cloud Computing: Who Wins?	http://itmanagement.earthweb.com/columns/article.php/3798591/Five-Companies-Shaping-Cloud-Computing-Who-Wins.htm
Cloud Computing Firms: Yes to Open Standards	http://itmanagement.earthweb.com/netsys/article.php/3830701/Cloud-Computing-Firms-Yes-to-Open-Standards.htm
Who Owns Data in the Cloud?	http://itmanagement.earthweb.com/netsys/article.php/3830896/Will-the-Kindle-Crisis-Kill-Cloud-Computing.htm
Software as a Service Model	http://bwhconsulting.com/Documents/BWHNewsletterV6-1.pdf
What Is Cloud Computing? (on YouTube)	http://www.youtube.com/watch?v=6PNuQHUiV3Q
Legal Thoughts on Cloud Computing	http://www.slideshare.net/brianharley/mason-hayescurran-philip-nolan-cloud-computing
Major Cloud Vendor Links	http://aws.amazon.com/ec2/ http://www.ibm.com/ibm/cloud/ http://www.salesforce.com/platform/ http://www.microsoft.com/azure/default.mspx http://code.google.com/appengine/whyappengine.html
Cloud Cost Links	http://calculator.s3.amazonaws.com/calc5.html http://www.ebizq.net/blogs/enterprise/2009/08/what_does_cloud_computing_actu.php#c6228 http://www.uptimesoftware.com/uptimeblog/2009/01/cost-of-cloud-computing-expensive/ http://www.infosysblogs.com/cloudcomputing/2009/07/is_price_the_only_criteria_to_2.html
Open Group Service Integration Maturity Model	http://www.opengroup.org/projects/osimm/
Cloud Security Alliance	http://www.darkreading.com/securityservices/security/government/showArticle.jhtml?articleID=221600333

9. DOCUMENT HISTORY

Version History

Ver.	Date	Sections	Description of Change
1.0	2009-12-12	All	Initial Document Release

10. GLOSSARY

Term	Definition
ASP	Application service provider – the more traditional hosted form of providing applications as a service
ARTS SOA Best Practices	Governing document from ARTS for all SOA artifacts to be published by ARTS
ARTS SOA Blueprint	Publication from ARTS on SOA for retailers
Mashing up services	Defined in this context as using multiple cloud applications for a data processing solution
Multitenant	Multiple organizations or departments within an organization accessing services provided by a cloud platform (typically cloud computing resources can be shared among multiple groups of consumers with significantly different requirements)
Open cloud	Provides services that are portable, interoperable, and manageable between multiple cloud providers
OSIMM	Open Group Service Integration Maturity Model – usually used to gauge the maturity of SOA adoption in a given organization
PCI compliance	Payment card industry compliance – required to ensure adequate security for storage of credit/debit card information and transactions
SLA	Service level agreement – a key component of services that acts as a contract between the service provider and service consumer
SNIA	Storage Networking Industry Association
Virtualization	Process of masking physical resources and making them available as logical units (combining many physical resources into one logical unit, splitting a single physical resource into multiple logical units, or any combination thereof), all in an attempt to make usage simpler and more efficient

