Designing Private and Hybrid Clouds

Architectural Best Practices

Brian Adler
Professional Services Architect
RightScale, Inc.
Abstract

The continued expansion of cloud adoption across a broad spectrum of industries and use cases has helped to solidify the establishment of Infrastructure-as-a-Service (IaaS) as a viable, cost-effective, and scalable IT delivery model. As the adoption of public cloud resources has proven effective across a diverse set of use cases, organizations have begun looking inward to find ways to more effectively and efficiently use their existing compute, storage, and networking resources in a similar model. Need breeds options in any industry, and cloud computing is no exception. This desire for internal resource utilization has given rise to numerous private and hybrid cloud technologies that provide tools for on-demand provisioning of compute, storage, and networking resources above and beyond what was available previously in both classic and virtualized datacenters.

RightScale, in particular its Professional Services group, has extensive experience helping customers in a variety of industries build private and hybrid clouds and then helping them manage those private and public resources. This white paper describes the considerations that you should make when embarking on the task of creating a private cloud from internal resource pools. It also explains the challenges that you will face at different branches of the decision tree. Additionally, this white paper explores the use cases for the allocation and utilization of these internal resources, focusing in particular on the technological considerations that need to be addressed in the implementation of these use cases. And lastly, it provides a reference architecture for private and hybrid clouds that is the framework for the customized architectures used by many RightScale customers.
In the not-too-distant past, major challenges facing CIOs and their IT teams were issues regarding power, space, and cooling of corporate datacenters. As the capabilities of server hardware continued to increase (Moore’s Law continues to stand the test of time), the workload generated by a typical application could no longer completely consume the resources of a single server. The numbers vary depending on the data source you refer to, but CPU utilization of standalone servers trended between an average of five to 20 percent. Regardless of where the true value may lie in this range, vast amounts of valuable resources were sitting idle, and burning capital with their power, space, and cooling requirements. Datacenter virtualization stepped in and helped to solve many of these problems.

With the implementation of multiple virtual machines (VMs) running within the same physical host, the classic setup of one server per physical host was expanded to a many-to-one relationship, providing huge cost savings. But once that vast improvement to the classic one-to-one design had been found, the issues regarding the procurement and provisioning of those resources still remained.

CIOs and their IT teams could compress more computing resources into the same physical and power footprint, but getting new VMs provisioned could still be a time-consuming and labor-intensive process. Typically, a support ticket would be filed with an IT department, which often resulted in separate tickets for the compute, network, and storage administrators. Manual action was then required on those tickets to provision the resource.

By creating a private cloud, you effectively eliminate the provisioning bottleneck present in the datacenter virtualization model while preserving cost savings as the same many-to-one relationship exists between VMs and physical hosts. By automating the procurement and provisioning of VMs, you can more efficiently provide these resources to any or all members of your organization.
Making the Move to a Private Cloud: Key Considerations

Many organizations with underutilized datacenter capacity are attracted to private and hybrid cloud implementations for the increased performance, better security controls, and greater cost efficiency that these environments can provide. But as with any new technology, it is important to consider the pros and cons of private and hybrid clouds and select the appropriate environment accordingly.

There are several areas of consideration that should guide your evaluation of private and hybrid clouds for your organization’s needs:

Workload and Infrastructure Interaction
Choose the correct workload for the infrastructure, or tailor the infrastructure to a specific workload. If the hardware allocated for the private cloud contains high-performance CPUs, then a processor-intensive application would fit nicely in the confines of the available resources. If high-bandwidth I/O is required, and the disk drives in use are not up to the task, you may be required to update the storage systems to meet the demands of the application.

Security
A pure private cloud configuration is attractive to organizations that have a known application workload footprint in terms of CPU, memory, and storage utilization, and that have potentially stringent security considerations. In the private cloud environment, all application traffic (and therefore all data) is contained locally within the organization, which can be of great benefit (if not required) for security compliance.

Latency
Network latencies are generally significantly reduced in a private cloud because traversal of the public Internet to the end user is typically not required. Storage performance can also be optimized in a private cloud environment with the use of high-performance storage hardware and the configuration of high-bandwidth connectivity to those devices.

Cloud Definitions

Because numerous definitions exist in the industry for terms specific to public, private, and hybrid clouds, these terms are defined in this white paper as follows:

**Private cloud:** A collection of compute, storage, and network resources for a single tenant that are accessed programmatically via an Application Programming Interface (API) endpoint.

**Public cloud:** A similar set of resources that is multi-tenant and is provided by a cloud vendor with access via an API endpoint.

**Multi-cloud:** An environment that spans two or more separate clouds, be they both public, both private, or one (or more) of each.

**Hybrid cloud:** An environment that spans one or more public clouds as well as one or more private clouds.
User Experience
For internal applications in particular, the geographic proximity of private cloud infrastructure to the consumers of the resources the private cloud provides is typically closer than the proximity of a public cloud environment. This further reduces latency and improves application performance and the resulting user experience.

Cost
A private cloud environment is in most cases less expensive on a VM-per-hour (operational expense) basis compared to a public cloud, and storage and network costs will most likely be significantly reduced. Data ingress to a public cloud may be free, but egress is typically charged on a sliding scale based on total bandwidth consumption. As such, a public cloud can prove prohibitively expensive for high-bandwidth applications.

1.2 Hybrid Clouds As a Strategic Option
The pure private cloud environment cannot always provide all the resources required by an application with unpredictable growth patterns. As an application grows in popularity and complexity, the available resources of the private cloud may no longer be able to support the growing user base. As a result, many organizations consider a hybrid cloud solution, which allows strategic access to both the private and public cloud resource pools.

A hybrid cloud architecture opens the application to the “infinite” resources of the public cloud. This spanning of multiple resource pools creates a unique set of challenges that need to be addressed, which include but are not limited to: security, latency, cost, and application complexity. Section 3 goes into greater detail on the technical considerations of hybrid cloud architectures and the specific hurdles that need to be overcome in these environments.

Many factors play into the decision of selecting the appropriate infrastructure environment for the desired workload. Private cloud infrastructure typically provides a more controlled and optimized environment for deploying application workloads, while allowing for tighter security and lower latencies. However, scalability can become an issue with a private cloud because resources are limited and finite. If internal resources are exhausted, an alternative (or supplemental) approach is required. Public clouds provide virtually infinite resources and provide an environment where applications can scale (theoretically) without bound. These environments are typically more costly at scale, have increased latencies, and can be complex when designing security into the solution.

Hybrid solutions are often declared to be “the best of both worlds.” And although at face value this may appear to be accurate, the spanning of multiple resource pools adds a new set of challenges, which while surmountable, are not trivial. Many RightScale customers come to the cloud with the concept of “cloud bursting” in mind. That is, the ability to use the public cloud to horizontally scale their server tiers when the capacity of their private cloud is exhausted. However, the complexities and costs of such a solution are often deterrents to a hybrid cloud implementation. These complexities are discussed in Sections 2 and 3.

A common use case for the hybrid cloud is for applications that have stringent security requirements (typically databases and/or shared storage systems), which are best placed in the private cloud infrastructure. The remaining components of these applications, which may have high availability requirements for front-end tiers such as load balancers and web and application servers, are best placed in the public cloud where they can scale as required. Although the security of in-flight data in these solutions needs to be addressed, the data at rest is within the confines of the private cloud infrastructure. Additional use cases for private and hybrid environments are covered in Section 2.

All cloud environments, whether they be public, private, multi-cloud or hybrid, also introduce management and provisioning challenges. Facilitating the deployment of application components across multiple resource pools in a coordinated manner is complicated due to differing APIs,
different implementations of infrastructure components (for example, attachable volumes versus local storage or SAN environments), and different server “sizes” and capabilities. A comprehensive cloud management solution can ease this burden by providing an orchestration layer that abstracts the API differences and disparate infrastructure components, enabling you to manage all your resources via a single user interface and API.

In subsequent sections of this white paper, some common use cases for private and hybrid cloud environments are described in more detail, as are the hardware, software and design considerations that come into play when following the best practices for developing private and hybrid clouds.
2
Reference Architectures and Use Cases

2.1
Reference Architectures

This section covers three separate customer use cases (A: Self-Service IT Portal, B: Scalable Applications with Uncertain Demand, and C: Disaster Recovery). Each of the applications described in these use cases is based on an architecture similar to one of the reference architectures presented in the figures depicted in this section. Use case A was implemented as a standalone private cloud, while both use cases B and C combined a private cloud and public cloud into a hybrid cloud solution.

One of the major components of any private cloud implementation (be it standalone or part of a hybrid cloud solution) is the software that provides access to the underlying physical resources. This software provides orchestration controls for the deployment of the virtual machines, storage offerings, and network configurations that are leveraged by the workloads of the private cloud. The product category for the vendors of these solutions is known by various names, the most prevalent being cloud infrastructure software.

Each of the current cloud infrastructure software providers uses varying nomenclature for its architectural components and allows for different configurations and setups. However, in the most basic case, the reference architectures of each of these options look very similar. Figure 1 shows a common architecture that is often used for proof-of-concept configurations but that is not recommended for production purposes due to the lack of redundancy within the architecture.

The end user accesses the cloud resources via an API call that is processed by the API endpoint. This API endpoint functions as the cloud management server/cloud storage.
controller, and coordinates the allocation of private cloud resources based on the API command executed. Each compute node of the resource pool has “local” storage, which can be either true local storage that is physically attached to the node, or that is accessed via a shared pool of storage over NFS, iSCSI, Fibre Channel, or similar mechanisms.

Each of the compute nodes hosts one or more VMs, with the underlying hypervisor managing access to the node’s resources by each of the resident VMs (bare-metal configurations without a hypervisor are also possible with some cloud infrastructure software offerings). The architecture shown in Figure 1 can be expanded to include multiple clusters that add both capacity to the solution as well as redundancy that can be used to increase the overall availability of the infrastructure. Additionally, redundant API endpoints can be added behind a set of load balancers to eliminate the single point of failure presented by a lone API endpoint, which increases the availability of the private cloud resource pool. Figure 2 depicts a generalized reference architecture for this configuration.

2.2

Use Case A: Self-Service IT Portal

Many RightScale customers use their private clouds to implement Self-Service IT portals. A RightScale customer in the pharmaceutical industry with a sizable market share built an on-premise private cloud by repurposing existing outdated and underutilized hardware resources. Several of the company’s business units required rapid access to compute and storage resources for short-term development and test projects, as well as for data analytics. Many of the workloads required by these business units were standardized, and could be parameterized via a small set of inputs configured at boot time.

Using RightScale, the IT team created a small set of ServerTemplates™ (a configuration “blueprint” for a server running on a cloud) and made these templates accessible
Four Considerations When Choosing a Private Cloud

1. **Workload and Infrastructure Interaction:** Choose the correct workload for the infrastructure, or tailor the infrastructure to a specific workload. For example, if the hardware allocated for your private cloud contains high-performance CPUs, then a processor-intensive application would be compatible within the confines of the available resources. If high-bandwidth I/O is required, and the disk drives in use are not up to the task, then you may be required to update the storage systems to meet the demands of the application.

2. **Security:** If you have a known application workload footprint in terms of CPU, memory, and storage utilization and have stringent security considerations, then a pure private cloud configuration may be your preferred choice. In the private cloud environment, all application traffic (and therefore all data) is contained locally within the organization, which can meet your security compliance needs.

3. **Latency:** You can generally achieve significantly reduced network latencies in a private cloud because traversal of the public Internet to the end user is typically not required. And you can optimize storage performance in a private cloud environment with the use of high-performance storage hardware and the configuration of high-bandwidth connectivity to those devices.

4. **Cost:** A private cloud environment is in most cases less expensive on a VM-per-hour (operational expense) basis compared to a public cloud, and storage and network costs will most likely be significantly reduced. Data ingress to a public cloud may be free, but egress is typically charged on a sliding scale based on total bandwidth consumption. Because of this, a public cloud can prove prohibitively expensive for high-bandwidth applications.

... to the business units’ end users. The end users logged in to the RightScale Dashboard and selected one of these preconfigured ServerTemplates. In a matter of minutes, they provisioned a fully configured and available server for their test and development purposes. Because each of these server instances was independent, an end user could perform whatever actions he needed without being concerned about modifying the environment of another end user.

When development and testing was completed, the end user terminated the server instance, which reduced operational costs and freed up resources that could be repurposed for another user. The business unit then launched the completed application in the public cloud due to the fact that the private cloud did not have sufficient resources for the anticipated user traffic. What is significant is that no modifications were needed because the ServerTemplates were created for multi-cloud deployments, enabling the workload to be run in either a public cloud or private cloud.

An important point to note is that the application in its entirety was encapsulated in a single resource pool – either the public cloud or the private cloud but not both simultaneously. As mentioned previously, many RightScale customers initially envision a “cloud bursting” scenario, in which they use the private cloud for the launch of their applications while intending to use the public cloud for additional capacity when their private cloud resources are exhausted. However, there are many complexities to the bursting scenario, which include (but are not limited to), cost, latency, and security.

In a true cloud-bursting scenario, you would use two unrelated resource pools, which implies the traversal of the public Internet from the application servers (which is most common) in the public cloud to the database servers in the private cloud. This cross-cloud traffic is subject to all the potential delays associated with the public Internet, and as such, latency targets cannot be assured. Similarly, the traffic needs to be secured as it is on the open Internet, and this adds complexity (and additional latency due to processing overhead) to the implementation.
And lastly, most cloud providers charge for egress from (and possible ingress to) their environment, which can add significantly to overall operational costs. The combination of public and private clouds is common in many hybrid cloud implementations in that both resource pools are used within the organization, but rarely within a single application.

2.3

Use Case B: Scalable Applications with Uncertain Demand

Contrary to the previous use case, a private cloud can also be used as a stable production environment, and the public cloud can be leveraged as a trial-and-error or application-proving-ground environment, similar to the hybrid cloud architecture shown in Figure 3. Another RightScale customer, a leader in social gaming, has built out a very sizable private cloud infrastructure using multiple regions, multiple zones within those regions, redundant API endpoints, and cross-zone database implementations. The physical infrastructure used in this private cloud comprises high-capacity commodity hardware so that the company can make more efficient use of its VMs. (For example, the work of two or three VMs in the public cloud can be handled by a single VM in the private cloud.)

In the early phases of the company’s private cloud development, it launched new applications in the public cloud because user adoption could not be accurately predicted. If the application was not well-received, it could run its course on the public cloud, and resources could be decommissioned as user activity diminished. If the application became popular quickly, the theoretically infinite resources of the public cloud could be leveraged. Ultimately, a stable workload for the application became apparent, and the company moved the application from the public cloud to the private cloud using a sophisticated set of migration tools.

Eventually the company greatly increased the capacity of its private cloud, which allowed for new applications to be...
launched in the private cloud environment, while its legacy applications largely remained in the public cloud. Over time, the company observed usage patterns and trends in the compute and networking demands of its applications in the public cloud and used this information to optimize its private cloud. This allowed for operational efficiencies that led to lower overall operational costs on a per-VM basis.

2.4

Use Case C: Disaster Recovery

Outages affecting IaaS vendors are rare, but they do occur. This reality has driven a growing number of RightScale customers to look for disaster recovery options. The majority of these companies do not have private cloud environments, so their primary resource pools are in the public cloud. Their disaster recovery resources also reside in public clouds – either in a separate region of the same IaaS provider, or with a different IaaS provider entirely.

For companies running their applications in a private cloud, the public cloud is typically leveraged as the disaster recovery environment. RightScale recommends that its customers implement a “warm” disaster recovery configuration as depicted in Figure 4 in which the database is replicated to a different IaaS environment (potentially another private cloud, but most commonly a public cloud). All other servers (load balancers, application servers, caching servers) are not in an operational state, but with RightScale they are ready to be launched with a single command and/or click of a button. Hence the “warm” configuration designation – the database server is the only component running and consuming resources.

This warm disaster recovery configuration has proven to be an effective solution for many RightScale customers in that it minimizes cost (because only the database is consuming resources) while greatly reducing the time required to bring the entire configuration to an operational state. The security, latency, and cost of the cross-cloud database communication contribute to the complexity of this disaster recovery solution, but these are typically surmountable issues.

![Figure 4 – Hybrid Cloud Warm Disaster Recovery Scenario](image-url)
3
Best Practices for Design and Implementation

3.1 Design Considerations
When building a private cloud (be it for standalone use or as part of a hybrid cloud environment), there are many design considerations that need to be addressed before you start the implementation. These initial considerations include:

- Location of the physical hardware
- Availability and redundancy configuration
- Intended workloads and use cases

Location of Physical Hardware
The physical location of your private cloud is the first decision you need to make. Many RightScale customers decide to host their private clouds in their own on-premise datacenters, but others opt to use a hosted solution in which a managed service provider’s facilities host the physical infrastructure. The latter is typically a more costly configuration, but many organizations do not have the redundant power, cooling, and networking capabilities that a managed service provider can offer. An on-premise solution allows for tighter control and easier and more efficient access to the physical infrastructure, which provides opportunities for maintenance and addition/deletion of resources. For organizations with strict security requirements, an on-premise solution may be required, depending on the compliance regulations that the application is subject to.

Availability and Redundancy Configuration
Once you establish the location for your private cloud, you need to consider the implementation details as they relate to availability requirements. Although a single zone (datacenter) within a single region (cloud) behind a single API endpoint is the easiest and most straightforward implementation, it is not a best practice from an availability...
3.1 Similarly, the use of separate regions further enhances the availability of the private cloud resources, but the infrastructure requirements of such a configuration are typically beyond the means of most organizations due to the cost and complexity involved. Each region would reside behind its own set of redundant endpoints, and be located in (preferably) different geographical locations with isolated power and networking configurations. Regions typically have no direct connectivity and rely on the traversal of the public Internet for communication between the regions. RightScale has numerous customers using private clouds with redundant API endpoints and multiple zones, but only a small number of these implementations encompass multiple regions.

It should also be noted that some traditional hosting and managed service providers are beginning to build out their own clouds using the cloud infrastructure software and topological best practices that are outlined in this white paper. But these vendors’ solutions are more similar to public clouds rather than private clouds as they are defined and discussed here.

Intended Workloads and Use Cases
When deciding whether a private or hybrid cloud implementation is the right choice, you should evaluate your use case and workload requirements in concert with the design and layout of your physical infrastructure. Factors to consider include:

- If the private cloud will be depended upon for high availability, then single points of failure (a single API endpoint or all nodes tied to the same power source, to give just two examples) need to be eliminated.
- If the private cloud will be used for CPU-intensive workloads, then the hardware must be selected appropriately (see Section 3.2 for more detail).
- If high bandwidth and low latency between nodes will be required, then appropriate networking equipment will need to be used for all relevant interconnects.
- If the private cloud will be used as a test-bed or a proof-of-concept platform that will be restructured frequently, then agility and flexibility are paramount, which will affect both the topology that is configured as well as the hardware that is selected.
At this stage of the design process, you should also consider the automation of the deployment of resources. You can use monolithic standalone images for the deployment of your server instances. Or, if you have a cloud management solution that allows for boot-time dynamic server configuration, you can use basic “vanilla” operating system images, which save time because you do not need to create new images when you make changes to the application or deploy new applications on your private cloud.

3.2 Hardware Considerations

When selecting hardware components to be incorporated into a private cloud environment, you may have several decisions to make. These include:

- **Compute**: Commodity or high end/specialized
- **Networking**: Flat/basic or advanced (VLANs, etc.)
- **Storage**: Commodity or high-performance/SSD/etc.

**Compute**

If you are constructing a new private cloud with no pre-existing hardware, then a case can be made for using commodity hardware. It is more cost-effective, and you can easily add new capacity to your private cloud because identical nodes are readily available. Additionally, you can simplify the maintenance of your private cloud due to the ubiquitousness of these components. However, commodity hardware may not have the performance characteristics (in terms of CPU, memory, IOPs, bandwidth, for example) required by the workloads destined for your private cloud.

Your use cases should drive your hardware selection process when possible. Several RightScale customers are using standard off-the-shelf commodity hardware in their private and hybrid cloud implementations so they may quickly add capacity when needed (this extra capacity can be added via entire server racks as opposed to a node at a time if the proper components are selected and software pre-configuration tasks are undertaken).

In other cases (which is a more common situation in the early phases of a new private or hybrid cloud), existing hardware is in place, and the intent is to reuse and repurpose this hardware. Because most organizations’ datacenters evolve over time, and new hardware is introduced over a period of months and years, the resources available in these scenarios are typically very disparate, which can lead to complications in maintenance and upgrades of the cloud infrastructure.

**Networking**

Your selection of networking hardware may also be driven by the use case and workloads destined for the private and/or hybrid cloud. High-bandwidth, low-latency routes between nodes may require the acquisition of costly, high-end networking equipment. Different cloud infrastructure software offerings support various networking models, and this may influence your decisions on purchasing network equipment as well. Your desired network model may dictate (or at least influence) the networking components (switches, routers, firewalls, load balancers) and configuration that are required for your private cloud implementation.

Additionally, some cloud infrastructure software provides for the integration of network hardware components (load balancers in particular) within the cloud infrastructure, and the use of these components may require a particular network model and/or topology. Again, the use cases and requirements of the cloud under consideration will drive many characteristics of the private cloud, but you must always keep in mind the repercussions that these decisions may have on other aspects of the infrastructure components and design.

**Storage**

In addition to the compute and network hardware of the private cloud, the storage requirements will dictate which physical storage systems are used. The various cloud infrastructure software offerings provide for different storage requirements and capabilities, and you should consider these carefully when selecting an implementation.

All of these technologies support the concepts of local storage as well as attachable volumes, but there are
options for how these storage mechanisms are physically implemented. In this regard, you can build your private cloud to have any manner of storage performance levels. You may implement storage using commodity drives or you may use solid-state devices for increased performance. Connectivity to these drives may be accomplished via numerous technologies, such as NFS, iSCSI, and Fibre Channel. As such, depending on the needs of the application workloads, along with the intersection of the mechanisms supported by the cloud infrastructure software and the budget considerations of the project, a wide combination of storage architectures is possible in the implementation of your private cloud.

3.3

Software Considerations

Cloud Infrastructure Software
Several commercial and open source cloud-enablement software solutions are currently available in this continually evolving space. Selecting the appropriate software solution for your private cloud is one of your key decision points and influences many, if not all, of the decisions you will make regarding your private cloud design and implementation.

Commercial solutions typically come with an open source counterpart, which may have a limited set of features or capabilities, while the paid version most often is full featured and fully supported by the vendor. Eucalyptus Systems is one cloud infrastructure software vendor that offers its solutions in this manner. Citrix CloudStack previously did the same, but now offers all code as open source and provides extensive vendor support only with its commercial version. There are other pure open source solutions such as OpenStack, and you can choose to work with systems integrators that will implement, support, and maintain these open source solutions.

Each of these cloud infrastructure software solutions uses concepts common to public clouds (zones and regions) but
Designing Private and Hybrid Clouds: Architectural Best Practices

15

Additionally, a cloud management solution provides tools for the automatic provisioning and configuration of cloud resources, allowing applications to dynamically (and proactively) respond to demand and adjust the deployed resources accordingly. Automating without a cloud management solution is possible, but without the level of abstraction provided by the solution, any implementation that you create will be compatible only with your current cloud infrastructure software and will not be portable to other environments.

A cloud management solution also provides a system-wide governance tool, allowing for visibility and control of user access and security protocols, granular auditing and logging of system access and resources, and fine-grained usage and metering controls. All these governance features are provided via a single consolidated system view and can span the private/hybrid/public cloud environments.

3.4

Implementation Process

The design and implementation of a private cloud involves a series of decisions, including:

- Hardware Procurement
- Cloud Infrastructure Software
- Cloud Topology
- Build or Buy

Hardware Procurement

Depending on how far your organization has traveled down the private/hybrid cloud path, many decisions may have already been made. Or circumstances may have dictated particular constraints — the hardware already exists, for example, so selecting the ideal compute, storage, and network equipment for the cloud may not be possible. If the situation permits, one of your initial considerations should be the hardware you intend to use in the cloud. The hardware should be tailored to your workload requirements. For example, commodity hardware for general-purpose tasks may be adequate,
but high-performance components are a better fit for resource-intensive workloads.

**Cloud Infrastructure Software**
A second consideration, and one that may drive many future design decisions of your private cloud, is the capabilities of the cloud infrastructure software. These capabilities may weigh heavily in your selection process because the specific design requirements of your private/hybrid cloud may not be supported by all vendor solutions. Or your design requirements may be more complex to configure and support with one solution as compared to another. So you should thoroughly vet cloud infrastructure solutions by starting with the vendor documentation and third-party research.

**Cloud Topology**
Once the hardware is in place and you have chosen your cloud infrastructure software, you must address the topological design of the cloud. The combination of these two choices may dictate (or limit) the choices that are available in regard to how you design and configure your private cloud. The reference architectures shown in Section 2 are common in many basic and intermediate private cloud environments, but more advanced configurations may involve much more complex physical and logical implementations.

**Build or Buy**
The final decision is whether to create your private cloud yourself or to outsource the implementation work. Many organizations have the resources in-house to handle the work involved in the implementation, but others contract with third parties to handle these details. Numerous systems integrators are adding cloud-building services to their product catalogs. Some focus on pure consulting services, others do hands-on work with customer equipment, and still others rapidly deploy entire racks of preconfigured servers to an onsite or hosted private cloud environment.

### 3.5 Management Process
Once you’ve made the decision to implement a private or hybrid cloud, you should also consider whether a cloud management solution can reduce the complexity of managing your implementation and reduce ongoing operational costs while preserving your ability to work with the vendors you prefer. When evaluating a cloud management solution, there are several key criteria for you to consider:

**Compatibility with Leading Vendor Solutions**
Ensure that you can build your private or hybrid cloud with existing and future investments in mind, including hypervisors from Citrix, Microsoft, and VMware; private cloud distributions including CloudStack, Eucalyptus, and OpenStack; and public clouds such as Amazon Web Services and Rackspace.

**Control and Security in One View**
Make your applications and infrastructure manageable by monitoring and controlling user access, network security, keys, and credentials in a unified view. Easily roll up and drill down your resource usage and costs.

**On-Demand, Self-Service Provisioning**
Enable your lines of business and developers to access the resources they need when they need them. Automated provisioning built on dynamic configuration management makes for manageable and quick-to-deploy applications.

**Freedom to Focus on Your Applications**
Manage to the application layer rather than the machine image to achieve true resource elasticity and self-service. Maintain environments that automatically shift with demand and resource failures and easily replicate and update those environments.

Depending on such factors as use case and implementation size, not every private or hybrid cloud implementation may benefit from a cloud management solution. But the convenience of managing an entire implementation in one place, being able to automate routine tasks, and the ability to maintain choice of vendors is an attractive model for many agile enterprises.
Conclusion

The acceptance of the public cloud as a viable IT strategy has resulted in many organizations looking at ways to use a similar IaaS model with their existing datacenter assets. And the recent emergence and maturity of cloud infrastructure software has made creating a private and/or hybrid cloud a more appealing and achievable goal. The extensive experience of RightScale in the cloud management space and its diverse customer base across many industry verticals has given the company a unique perspective on the emerging private/hybrid cloud sector. The RightScale Professional Services group has assisted numerous companies with the installation, configuration, and management of their on-premise resource pools. Many of the initial private cloud projects were undertaken as a proving ground for the concept, but a growing number of companies are taking the next step and using their private clouds in production environments as standalone resource pools or as part of a larger hybrid cloud infrastructure spanning one or more public clouds.

In the last few years, the private cloud market was viewed as an interesting science or academic project. But with the maturity of cloud-enablement technologies, this market has grown much faster than many anticipated. In some industries, the use of a private or hybrid cloud is a more common choice than a pure public deployment. As the software and hardware offerings in this market continue to evolve, the opportunities for private and hybrid cloud adoption will continue to grow as additional workloads become viable options for these resource pools. And “niche” clouds (optimized for storage, or network latency, or auditable security environments, for example) will continue to emerge.

As with many complex technologies, there is no “one-size-fits-all” solution when it comes to determining the design and implementation of a private or hybrid cloud. But by considering your use case and defining your workloads in concert with the overall design of your architecture, and by evaluating the hardware and software technologies that best complement your needs, you can develop an architecture that supports your technological and business goals.