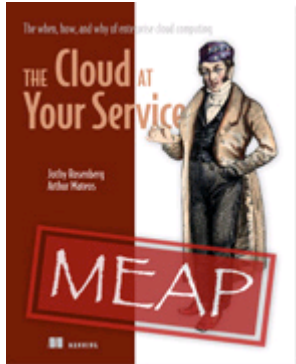


When are Private Clouds a Good Idea?

Excerpted from



[The Cloud at Your Service](#)

EARLY ACCESS EDITION

The when, how, and why of enterprise cloud computing
Jothy Rosenberg and Arthur Mateos

MEAP Release: February 2010

Softbound print: Summer 2010 | 200 pages

ISBN: 9781935182528

This article is taken from the book [The Cloud at Your Service](#). The authors define the term private cloud and discuss issues to consider before opting for private clouds and concerns about deploying a private cloud.

Get **35% off** any version of [The Cloud at Your Service](#), with the checkout code **fcc35**. Offer is only valid through <http://www.manning.com>.

Many vendors with products and services that involve virtualization, storage, or data center automation will claim that they have a private cloud offering.

Private cloud can be viewed as comprising a subset of the five main principles behind public clouds:

- Off premises: someone else owns the hardware assets
- Virtualization: high utilization of assets
- Elasticity: dynamic scale without capex
- Automation: build, deploy, configure, provision, and move—all without manual intervention
- Metered billing: per usage business model; pay for what you use.

For private clouds, the three principles associated with the technical attributes still hold true, namely *virtualization*, *elasticity*, and *automation* are technical attributes and translate directly from public clouds to private clouds. The other two relate more specifically to the business attributes of public clouds and are less directly applicable to the private cloud.

What is a private cloud?

Private clouds, by definition, are not off premises and are owned by the cloud operator. Metered billing may be implemented in certain cases for private clouds within large organizations but are not strictly necessary.

DEFINITION: PRIVATE CLOUD

Private cloud (also called internal cloud or corporate cloud) is the term for a computing architecture that provides hosted services to a specific group of people behind a firewall. A private cloud leverages virtualization, automation, and distributed computing to provide on-demand elastic computing capacity to internal users.

Public cloud providers are spending a great deal of money in new data centers to power their cloud initiatives, with Google investing roughly \$2.3 billion in 2008 for its buildout.¹

Public cloud spending

At first glance, in the face of this massive investment, it may seem like a foolhardy proposition to attempt to go it alone and build one's own private cloud. We should remember, however, that large IT organizations have a long history of providing data center services, many with longer track records of doing so than most of today's incumbent cloud providers (i.e., Amazon, Google, and Microsoft). There are tremendous amounts of resources and past investment in hardware and data center assets. Certainly these can be put to good use.

Over the last decade, many large enterprises have launched virtualization projects and initiatives and have reaped the benefit of increased resource utilization and efficiencies. Those that have done so are one step closer toward having a private cloud. This is one of the three key technology principles for private cloud computing. The only incremental changes needed to have a private cloud are the addition of elasticity and cloud automation technologies.

Things to consider before opting for private cloud

There are four primary considerations at play that drive the consideration of a private cloud as a deployment strategy summarized in Table 1.

The security and availability constraints of target applications and data and the degree to which they must be under direct control may dictate whether a public cloud option is even viable, or whether a private cloud solution needs to be considered. The size and needs of the user community should be sufficiently large and the organization should be large enough to be able to have economies of scale when purchasing capital equipment for a private cloud deployment to make sense.

Table 1 Private cloud considerations

Private Cloud Considerations	
Security	Applications that require direct control and custody over data for security or privacy reasons
Availability	Applications that require certainty in access to a defined set of computing resources that cannot be guaranteed in a shared resource pool environment
User Community	Organizations with large numbers of users, perhaps geographically distributed, that need access to utility computing resources
Economics of Scale	Existing data center and hardware resources that can be used and the ability to purchase capital equipment at favorable pricing levels

For a private cloud to be a good option, the organization's security and availability requirements should be high and the scope of the user base and the purchasing power of the organization must be sufficiently strong.

¹<http://www.datacenterknowledge.com/archives/2009/05/18/facebook-20-million-a-year-on-data-centers/>

Security considerations

There are some applications for which security considerations reasons don't allow public cloud computing. For example, many government organizations have applications that deal with confidential or top secret data that, under no circumstances, may be put at risk, such as those dealing with national security. There are other applications in other industries where regulatory requirements might make one think twice before deploying them to a public cloud.

Public cloud providers are aware that security is a main blocking or gating factor for many enterprises and have devoted significant resources to designing and proving their ability to deal with secure data. Amazon, for example, recently announced that it achieved SAS 70 Type II certification for AWS. A SAS 70 Type II audit is a rigorous procedure that ensures the provider has the appropriate processes and infrastructure in place to handle data securely and with high availability for its customers. Amazon has also made claims that its infrastructure has been designed so that it can support the requirements of regulatory frameworks such as the Health Insurance Portability and Accounting Act (HIPAA). HIPAA spells out the measures that organizations in healthcare industry must adhere to in order to ensure the privacy of their patient's data. Having the hooks to enable HIPAA compliance and implementing an application that is HIPAA compliant are two very different things. Best practices must be developed and experience gained in implementing HIPAA compliant applications in the public cloud before most enterprises will be comfortable with this mode of deployment.

Certainty of resource availability

Although we think of the cloud as an infinite resource pool from which resources can be drawn, this is not always the case. As an example of this, consider an application that requires a huge number of resources for doing massive amounts of processing in a short time window. As of late 2009, Amazon has advised its users that it could not guarantee the availability of 500 XL instances (where XL instance are high-compute resources with 8 CPU virtuals) at any given time from a specific availability zone. For cases in which resources in excess of 1000 XL instances were required, Amazon requested that they be given a week's notice to improve the chances of the available resources.

Resource constraints are a much more serious matter in smaller cloud deployments. Rackspace, for example, in late 2009 imposed a limit of 50 virtual instances in their environment per day for any given user. Overall, the total capacity in these systems should improve going forward but even then, there is still the caveat related to variations in demand caused by overlapping requirements from multiple different public cloud customers. Electric utilities, which have been running now for over a century still run into capacity issues in the heat of the summer when the demand for electricity to power air conditioning can cause brownouts due to a mismatch of available supply and demand. One could easily imagine the same thing happening in the cloud context if all e-commerce sites were public cloud consumers and witnessed 10x traffic spikes on Black Friday. Perhaps at some point, cloud pricing would take into account the variability in demand and introduce variable pricing.

Large utility computing community

For a company with relatively few users that require utility computing resources, a good virtualized infrastructure would probably suffice. If, however, there are many constituents within the organization that can take advantage of a generalized infrastructure for their needs, then the added complexity and sophistication of a cloud infrastructure may make sense. By implementing a private cloud, you introduce the concept of multi-tenancy and, hence, the ability to segment and isolate individual groups and users.

Economies of scale

There are two potential primary advantages that a public cloud provider has from an economic perspective over a company interested in running their own private cloud. The first relates to the physical resources required to run a cloud. The second aspect relates to the expertise required to run and maintain a cloud infrastructure. Because they have been designing for scale, public cloud providers have been creating infrastructure where one of the primary objectives is reducing the number of resources required to operate a data center. In most cases, conventional IT organizations require more engineers and technicians to run a smaller data center. By migrating to a cloud style

deployment, it is possible that they will save money over their existing deployment. This might, however, require a retooling of their current resources or hiring a smaller number of more skilled resources.

Some concerns about deploying a private cloud

Before you or anyone jumps to deploying a private cloud, you should assess four major concerns about doing so.

Private clouds are small scale

There's a reason why most innovative cloud computing providers have their roots in powering consumer web technology—that's where the numbers are. Very few corporate data centers will see anything close to the type of volume seen by these vendors. And, volume drives cost through the huge economies of scale we have discussed.

Legacy application don't cloudify easily

Legacy applications moved to a private cloud will see marginal improvements at best. There's only so much you can achieve without truly rearchitecting these applications to a cloud infrastructure.

On-premise does not necessarily mean more secure

The biggest driver towards private clouds has been fear, uncertainty, and doubt about security. For many, it just feels more secure to have your data behind your firewall in a data center that you control. But, unless your company spends more money and energy thinking about security than Amazon, Google, and Salesforce, that is not true.

Do what you do best

There's no simple set of tricks that an operator of a data center can borrow from Amazon or Google. These companies make their living operating the world's largest data centers. They are constantly optimizing how they operate based on real-time performance feedback from millions of transactions. While you can try to learn from and emulate them, your rate of innovation will never be the same—private Clouds will always be many, many steps behind the public Clouds.

Summary

Private cloud computing is a potential alternative deployment option that is available and may make sense for very large enterprises and organizations. For organizations with enough scale, buying power, and expertise, private clouds offer the advantages of increased control, predictability, and security. There are many options available including building one from open source technologies, utilizing proprietary purpose built solutions, or by partnering with service providers willing to allocate or partition dedicated resources for a private cloud.