

VMware vSphere, the First Cloud Operating System, Provides an Evolutionary, Non-disruptive Path to Cloud Computing



IT Is the Business

In today's ubiquitously networked, highly automated, technology-enabled economy, business capability has become inextricably linked to IT capability. The hard truth, however, is that IT inhibits as much as it enables the business. This seeming contradiction is explained by the evolution of IT during the past 50 years since the invention of the integrated circuit. This evolution has been characterized by a tension between forward momentum and the inertia of legacy systems. On the one hand, there is constant break-neck innovation and the pursuit of the "next big thing." On the other hand, nothing, really, goes away, and each new generation of computing is layered on top of the old. Walk into any datacenter, and you will find an old mainframe application chugging away – probably with web portals and Java-based front ends that were bolted on during the dot-com boom of the 1990s. This pile up of geological layers of technologies in the datacenter has resulted in overwhelming complexity.

The long view of technology shows a sweeping panorama of Moore's law and amazing human creativity plunging forward at mind-boggling speed. But zoom in and you see a story of incrementalism. Datacenters today are hodgepodes of architectural, technology, vendor, and management decisions made incrementally over time. And each decision tends to get hard-coded into the infrastructure. In the long haul, however, the unintended consequences of incrementalism and hard-coded complexity add up.

So this is the dark side of the IT story: Increasingly layered, complex, hard-coded, and sprawling, IT over time has grown inflexible, brittle and costly. Today, as a result, the typical company spends more than 40 percent of its IT budget on infrastructure maintenance and around 30 percent on application maintenance. That adds up to 70 – 80 percent of IT spending simply to maintaining the status-quo. Think of this as a complexity tax. And that is one steep tax. Imagine being taxed 70 percent on your income – you would have a hard time making a living.

In addition to devouring the lion's share of the budget, the complexity of IT exacts an even greater cost in terms of lost opportunity. Rather than innovating, IT organizations spend most of their time fighting fires, fixing problems with the IT "plumbing", jury-rigging the old stuff to run with the new stuff, and running routine maintenance tasks just to keep the lights on in the datacenter. With less than 30 percent of the IT budget available for innovation and competitive advantage, business is starved for new and updated capabilities.

We have reached a point where business opportunity is literally trapped in IT infrastructure. Clearly, the traditional model of IT has become inoperable.

Cloud Computing: A Solution on the Horizon

This problem – this complexity – is the fuel driving the rise of virtualization over the past ten years and the current buzz about cloud computing. As the chart from Google Trends shows below, while cloud computing wasn't even on the radar before August 2007, public interest has been spiking since then with a vengeance.

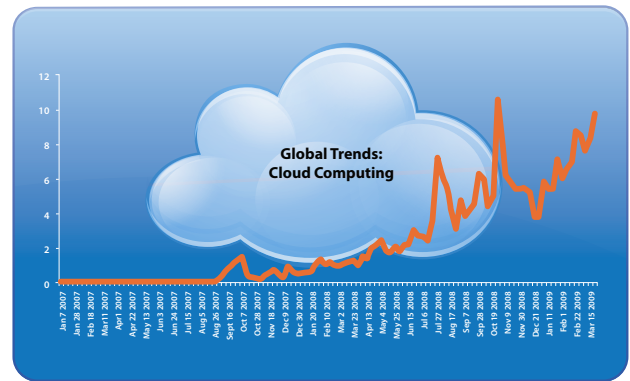


Figure 1. Google Trends: Cloud Computing, Jan 2007 - Mar 2009*

It's no accident that we paired virtualization and cloud computing in the sentence above, and in the order in which they're stated. Virtualization is a crucial enabler for cloud computing.

Cloud computing is a new style of computing that makes IT infrastructure available as a ubiquitous, easily accessible, and reliable utility service, conceptually similar to the telephone or electricity. Of course, the IT industry has been moving along this trendline for years, with software-as-a-service (and platform-as-a-service, infrastructure-as-a-service, hardware-as-a-service, etc.). At last, the reality is catching up to the dream. The idea of cloud computing is to transform IT from assets that you own into services that you use and pay for on a per-use basis. With electricity, for example, it's not necessary for each of us to have our own generator; we simply plug our appliances into the wall, turn them on, and voilà. We don't need to understand how or where the electricity is generated. We don't need to be able to read a wiring diagram. We only need to pay the monthly utility bills to keep the current flowing. As for technical knowledge, if we can screw in a light bulb and push a plug into a socket, we're all set.

* Note; The numbers on the left of the graph do not directly reflect numbers of searches, but are scaled and normalized by Google to reflect relative volume trends, as explained here: <http://www.google.com/support/insights/bin/answer.py?answer=87285&cbid=11utufgy0t0l&src=cb&lev=answer>.

Cloud computing transforms IT into a service that is:

- Ubiquitously available
- Easy to access
- Reliable
- Affordable – pay as you go, for only

But There Are Challenges for the Enterprise

That's the good news about cloud computing. The not-so-good news is that some implementations of cloud computing require businesses to rewrite or customize their applications so that they can work in an external, public cloud infrastructure.

Early industry conversations about cloud computing have focused on the idea of external, public cloud infrastructures such as Amazon EC2 or SaaS offerings such as Google Apps. From the standpoint of the enterprise, however, these external cloud services still represent niche offerings for one simple reason: to put your apps on these clouds requires re-writing the applications to some extent. If you use Google, for example, you have to use Google's datastore API to persist state in your apps; if you use Amazon, you've got to use EC2's database API and virtual file format (called the Amazon Machine Image).

This proprietary cloud model is not a viable option for most businesses for three key reasons:

- 1. Disruption.** Businesses can't afford the time, money, and disruption it would take to rewrite their applications to the new APIs required by proprietary cloud service providers.
- 2. Lock in.** Re-writing applications for proprietary cloud architectures will lead to the same sort of inflexibility that enterprises face today in their own datacenters. It's that "Hotel California" syndrome the Eagles sang about: You can check in, but you can't check out.
- 3. Inadequate service levels.** Current offerings are not, in general, enterprise-ready. The vendors are not yet able to commit to the service level agreements (SLAs) for availability and security that companies require in order to entrust their core applications to the public cloud.

Still, the Potential Benefits of the Right Solution Are Huge

Yet, despite these caveats, the excitement over cloud computing continues to grow in the business world. It's being driven by the promise of enormous savings that result from the paradigm shift: from IT as a physical asset – and the capital costs that go with it – to IT as a service that you pay for as you use it.

Consider the case of the Preferred Hotel Group, which has moved its datacenter operations to Terremark's cloud computing platform. For 10 virtual servers – seven full-time and

three ready to provide additional capacity to meet spikes in demand – the chain of luxury hotels is paying \$16,000 a month. "Everyone is checking their budgets now," says Chad Swartz, senior manager of IT operations for the group. "If you go to the board, is it an easier sell to say we need to spend \$200,000 in capital costs and \$10,000 a month? Or just pay a \$10,000 implementation cost and \$16,000 per month? The cloud environment is going to explode, if just for the cost savings."

Enterprises want the benefits of cloud computing – but they want these benefits on their own terms:

- They want a solution that is non-disruptive, enabling businesses to use their existing applications without having to re-write them, and at the same time accommodating new application architectures as they're invented. In other words, they want a practical, painless path to a cloud computing infrastructure that maximizes application choice and is future-proof.
- They want a solution that is enterprise-ready and provides uncompromised control over the service levels they're responsible for providing to the business.
- They want a solution that allows choice and avoids lock-in to any specific application, operating system or hardware architecture.

Transforming the Datacenter to a Private Cloud

There is a solution that meets these requirements. The solution is to bring the cloud to the enterprise, rather than forcing the enterprise to port its entire IT infrastructure to a proprietary external cloud. The solution, in other words, is to transform the internal datacenter into a private cloud that, optionally, can federate with external cloud services in a secure way.

VMware® believes that the realistic road to cloud computing for businesses is to enable the internal datacenter to act with the characteristics of a cloud. With such a solution, your IT group can provision services and capacity to users on a self-service basis, automate management tasks, and chargeback billing to the business. Your technical staff can become a team of service providers rather than an emergency fix-it squad, maintaining complete control over the company's IT and meeting required SLAs with precision. They don't have to rewrite applications but can use them as is. And your internal IT consumers do not have to understand how the technology works or what lies behind the service; they just plug in and voilà (and then pay for what they use via a chargeback service).

And, importantly, this is not a gut-rehab of your datacenter. It's an evolutionary journey – with revolutionary benefits. And, if you've implemented virtualization to any degree, you've already taken the first critical steps.

Many companies started on the path to cloud computing by consolidating servers through virtualization, in order to get a handle on server sprawl. A typical case – among many tens of thousands – is Independence Blue Cross, the largest health insurer in Philadelphia with 3.4 million members. “We needed a cost-effective solution that would manage server sprawl, and ensure we had a high-performance application environment that could handle our business as it grew,” says Michael Garber, Director of Distributed Infrastructure at Independence Blue Cross. “VMware met these needs. In fact, the performance of many of our Microsoft applications has actually improved since we started running them on the VMware platform. The product has paid for itself in less than 16 months and helped us avoid more than \$1 million in hardware costs. We now have 386 virtual machines running on 48 physical hosts. We’ve increased our processor utilization from 5 percent to 75 percent, and we still have extra horsepower to add more virtual machines.” And in so doing, the company also has moved well along the path to the cloud.

The Private Cloud: Bridging Internal and External Clouds

Here, to avoid confusion, a definition is in order. There are lots of competing definitions for aspects of cloud computing flying around. One is the “private cloud”. By this, some mean simply an internal cloud. And a public cloud infrastructure, such as Amazon or Google, is dubbed an external cloud.

VMware defines a private cloud – what some call a “hybrid cloud” – as a secure computing environment that allows computing capacity from both internal and external clouds to interoperate and be delivered much like a utility. The analogy is a virtual private network (VPN), which uses the public network but is controlled (in terms of access and security) by a particular company.

A private cloud brings unprecedented levels of flexibility, control, efficiency, resiliency, and manageability to datacenters and allows any application – legacy, server-based, desktop, or those built on new application frameworks – to be delivered as a service. Three technology components are necessary to build this sort of private cloud, which are shown in Figure 2:

1. A cloud operating system
2. A new management model
3. Federation services to connect multiple datacenters (whether internal or external clouds) into one private cloud and open standards to ensure the easy porting of applications and data when “burst-up” capability (or on-demand capacity) is needed

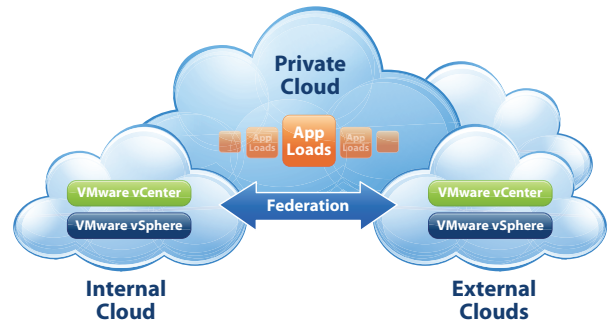


Figure 2. The Components of a private cloud spanning internal and external datacenters

Let’s examine each of these components in turn.

The Cloud Operating System

A cloud operating system is a new category of software that is specifically designed to holistically manage large collections of infrastructure – CPUs, storage, networking – as a seamless, flexible, and dynamic operating environment. Unlike traditional operating systems that manage an individual machine, the cloud operating system aggregates the infrastructure of an entire datacenter infrastructure to create a single, powerful “compute plant” whose resources can be allocated very quickly and dynamically to any application that needs the resources. Just as in the electric grid, where power from generation plants can be delivered quickly to any new appliance that is plugged in, the resources of this “compute plant” can be dynamically allocated to any newly provisioned application.

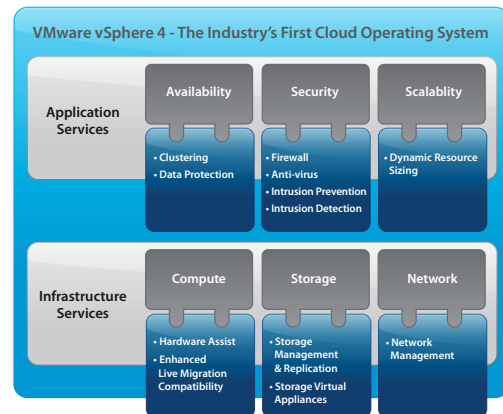


Figure 3. Core capabilities of the VMware vSphere cloud operating system

To create this “compute plant”, a cloud OS must provide these capabilities:

- **Infrastructure services** that comprehensively virtualize server, storage, and network resources, and aggregate them into a shared mainframe-like computing platform that is resilient and capable of running the most demanding applications with near-native performance.
- **Application services** that provide built-in service-level controls to all applications.

In April 2009, VMware delivered the industry’s first cloud OS, VMware vSphere™ 4. VMware vSphere meets these requirements with the following capabilities:

Infrastructure services

- **VMware vCompute** services virtualize server resources and aggregate these resources into logical pools that can be precisely allocated to applications. VMware vCompute services include VMware ESX™, the industry’s leading hypervisor; as well as unique capabilities to automatically allocate resources across virtual machines and clusters; and capabilities to maximize energy efficiency by continuously optimizing server power consumption.
- **VMware vStorage** services abstract storage resources from the complexity of underlying hardware systems to enable the most efficient utilization of storage capacity in virtualized environments.
- **VMware vNetwork** services enable optimal administration and management of networking in virtual environments.

Application services

- **Availability services** enable IT administrators to provide applications with varying levels of high availability depending on their priority, without the need for complex redundant hardware or clustering software. Through such unique capabilities as VMware VMotion™ and VMware Storage VMotion, scheduled downtime due to server maintenance or storage maintenance can be eliminated; and unscheduled downtime or data loss can be eliminated or minimized through unique features such as VMware High Availability, VMware Fault Tolerance, and VMware Data Recovery.
- **Security services** allow IT to provide applications with the appropriate level of enforcement of security policies no matter where they run – a key to controlling security across internal and external clouds in a private cloud environment. VMware vShield Zones simplify application security by enforcing corporate security policies at the application level in a shared environment, while still maintaining trust and network segmentation of users and sensitive data. VMware VMsafe™ enables the use of security products that work in

conjunction with the virtualization layer to provide higher levels of security to virtual machines than even physical servers.

- **Scalability services** deliver the right amount of resources to every application non-disruptively, based on its need, and include load-balancing and “hot add”, plug, and extend capabilities for CPU, memory, virtual storage, network devices, and virtual disks.

VMware vCenter™ Server is the central management server for organizing, monitoring, and configuring an IT environment through a single interface.

A New Management Model for the Private Cloud

Delivering IT as a service in the private cloud – across internal and external clouds – requires a new management model.

This new model must focus on the properties of that service, rather than on the underlying infrastructure components. In other words, the private cloud requires a service level-driven management model that allows IT to manage what business users and application owners understand – the degree of availability and security the service needs – to enable business users and application owners to easily request services, specify desired service levels, and then consume those services without understanding the intricacies of the underlying infrastructure. At the same time, this new SLA-driven management model must also translate the humanly comprehensible language into machine instructions, so that IT can manage, dynamically and with a high degree of automation, the underlying resources to ensure the desired quality of service to the business.

VMware vCenter is a comprehensive set of solutions for managing IT as a service with highly-automated, policy-driven administration of operations from self-service request through chargeback. With VMware vCenter products, VMware uniquely offers a comprehensive set of management capabilities to automatically provision new virtual machines and applications, to ensure compliance with established configuration standards, to right-size every element in the environment, to predict and manage capacity, and to allocate the costs back as appropriate to the service consumers. These management capabilities include:

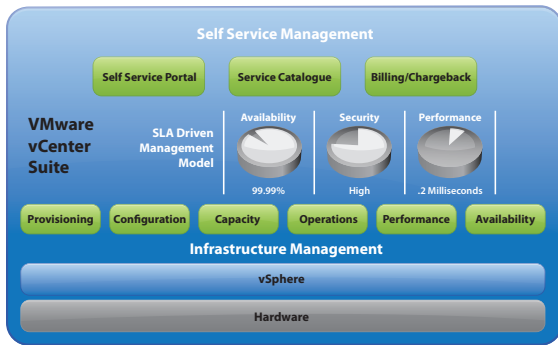


Figure 4. Infrastructure Management with VMware vCenter solutions

- **VMware vCenter CapacityIQ** helps ensure that adequate capacity is available to virtual machines, resource pools, and entire datacenters by modeling the effect of capacity changes, forecasting shortfalls, and proactively reclaiming excess capacity.
- **VMware vCenter AppSpeed** provides visibility into end-user experience, service level reporting, and proactive performance management for multi-tier applications, including virtualized and physical elements.
- **VMware vCenter ConfigControl** allows administrators to maintain visibility into and ensure compliance of configuration state in a dynamic virtual environment.
- **VMware vCenter Lifecycle Manager** enables administrators to implement a consistent, automated process for managing the lifecycle of virtual machines, from provisioning to retirement.

Self-Service Management

- **VMware vCenter Lab Manager** simplifies development and QA environments by giving users on-demand access to common system configurations, while administrators maintain control over policies and dramatically reduce server, storage, and provisioning costs.
- **VMware vCenter Stage Manager** helps to reduce server sprawl and the risk of configuration drift by enabling IT and application administrators to clone system configurations easily, and maintain consistent virtual machine images throughout the release process.
- **VMware vCenter Chargeback** helps administrators ensure accountability across the business by understanding, allocating, and reporting on costs associated with the use of virtual infrastructure.

As a whole, this comprehensive set of management capabilities gives businesses the choice of wholly owning and operating IT internally, or renting compute capacity from a cloud service provider using new technologies that federate internal and external clouds to provide seamless “flexing” of applications on-premise and off-premise, while guaranteeing the appropriate levels of security.

Moreover, VMware’s cloud management capabilities plug seamlessly into broader enterprise frameworks, allowing for a choice of management tools for non-x86 and x86 environments.

Federation Services to Connect Multiple Clouds in a Private Cloud

In addition to delivering the efficiency and flexibility of the cloud to internal datacenters, VMware is working in parallel with hosting and service providers to enable compatible external cloud infrastructures – infrastructures that do not require rewriting your existing applications and that are future-proofed to accommodate new application types as they are developed.

VMware vCloud™ is an industry initiative designed to enable federation between external and internal clouds to provide the elasticity for the private cloud. VMware is working in concert with more than 400 partners and service providers to achieve this goal, including industry leaders such as BT, EngineYard, IT Structures, Logica, Melbourne IT, Rackspace, Savvis, SunGard, Terremark (see the Preferred Hotel Group example above), T-Systems, and Verizon Business. These service providers are either currently offering or plan to offer VMware vCloud services with the security, service levels, and application compatibility required for enterprises to confidently incorporate them into their private clouds.

Conclusion

Cloud computing is not a decade away, not some distant vision. It's here today. It's not all-or-nothing. It's not a radical, unaffordably expensive, messy and disruptive gut-rehab of the datacenter. And it's not something that's only for the large global enterprise either.

Rather, it's a practical, non-disruptive evolutionary process; a path of virtualization adoption that is being led by VMware, where every step pays for itself with proven, powerful benefits. Customers that virtualize to reduce server sprawl and capital costs find that, at the same time, they have begun to pool IT resources and their datacenter is already more resilient, flexible, service-oriented – and more cloud-like.

With each step along the path, the datacenter becomes more cloud-like and applications become more service-like – faster, more efficient, more reliable, easier to manage. The end result is an internal cloud, poised to federate with external clouds as needed for additional capacity on demand.

Now, with the availability of VMware vSphere, the industry's first cloud operating system, the path to the next major evolutionary stage of IT – the private cloud – is clear. With VMware vSphere, a comprehensive set of management tools, and the VMware vCloud Initiative to ensure cloud compatibility, VMware enables IT organizations to abstract applications and information from the complexity of the underlying infrastructure – whether it's internal or external – and deliver applications as a service over which they have complete control.

Through federation and common management services between clouds, the VMware approach offers a seamless, dynamic operating environment that enables the bridging of internal resources with available external resources, helping businesses achieve the full flexibility and benefits of cloud computing. Today.

Finally, your IT organization can focus not on the overwhelming complexity of the datacenter, but on providing IT as a well-managed service. And they can spend their time, rather than putting out fires and wrestling with the plumbing, purely on the support and enablement of business value.



VMware, Inc. 3401 Hillview Ave Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com

Copyright © 2009 VMware, Inc. All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws. VMware products are covered by one or more patents listed at <http://www.vmware.com/go/patents>.

VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies. Item No: VMW_09Q2_WP_Cloud_OS_P8_R1

