WHITE PAPER

Open Source Cloud System Software

Sponsored by: The Linux Foundation

Gary Chen

November 2012

IDC OPINION

Cloud computing is changing the IT market in a multitude of ways. Public clouds are changing the way enterprises deploy computing resources, and private clouds are changing how enterprises approach their own IT. Open source software has played a leading role in the development of public clouds and now is moving into the private cloud space. Most open source cloud software builds on existing open source (Xen and KVM) or proprietary hypervisors and operating systems (Linux) by developing a new complementary class of software that IDC terms "cloud system software." Cloud system software provides abstraction and APIs at a higher level than a hypervisor. Open source cloud system software such as CloudStack, Eucalyptus, and OpenStack allows enterprises to build "Amazon style" clouds in their own datacenters. This IDC White Paper examines the results of a recent IDC enterprise cloud system software survey that probes the use of cloud system software in the enterprise and the role of open source and community.

METHODOLOGY

The Internet-based survey was conducted in August 2012 and was completed by 282 qualified participants in the United States. All of the respondents were enterprise IT customers who either had deployed a private cloud or were interested in deploying a private cloud. Respondents were IT management or administrators who make or influence selection and purchasing decisions for private cloud software in their organizations.

Respondents were distributed across market segments defined by company size. Small and medium-sized businesses (SMBs) with fewer than 500 employees were excluded from this survey.

Note: All the numbers in the document may not be exact due to rounding.

IN THIS WHITE PAPER

This IDC White Paper examines the results of a recent IDC enterprise cloud system software survey that probes the use of cloud system software in the enterprise and the role of open source and open cloud.

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

SITUATION OVERVIEW

Server virtualization was a catalyst in the market that ignited many of the datacenter and cloud trends we see today. Virtualization is the new default in datacenters, with virtual servers outshipping physical servers for the past two years. About two-thirds of all new workloads are deployed on virtual servers.

After VMware established the x86 server virtualization market, the open source community quickly began developing its own hypervisors. Xen was one of the first projects and was quickly supported by many major Linux distributions and formed the core of several commercial virtualization offerings. It also gained a high profile as part of large cloud installations, such as Amazon Web Services and the Rackspace Cloud. More recently, KVM was added to the Linux kernel, allowing Linux to become a hypervisor and run any unmodified x86 guest operating system. Other notable open source virtualization projects include oVirt, a management system for KVM, and LXC, a lightweight container virtualization technology that is also included with the Linux kernel.

On its own, virtualization brought consolidation and efficiency to servers and datacenters. However, it also had dramatic effects on the rest of the datacenter ecosystem, such as storage and networking, which had to transform themselves to work in virtualized environments. A holistic approach to virtualization is essential to the success of virtualization deployments today. As virtualization has increasingly become a crucial software component of the datacenter, it has also become the foundation of the emerging cloud market.

Cloud is often defined in many different ways, but at its essence, it is a delivery model for IT that provides elastic, agile services in an on-demand, pay-as-you-go fashion. However, a service model doesn't necessarily imply that a new architecture is needed. For example, a very small-scale cloud can be implemented using a traditional virtualized infrastructure with a Web portal on top. However, to accommodate more complex and large-scale deployments, many are beginning to use different architectures that are specifically designed for cloud.

While virtualization is generally considered a foundation for cloud, it is not always so. Many public software-as-a-service (SaaS) clouds today don't use virtualization, instead providing high levels of service and performance using clusters of bare-metal servers. This is made possible when an application's requirements are well-known and capacity increments are measured in whole servers or racks.

What really changed the way people thought about public clouds, from both a service delivery perspective and an architectural perspective, was Amazon Web Services. Amazon broadened virtualization beyond single servers to entire datacenters, in a highly automated and super-scale fashion. In this manner, Amazon provided much of the inspiration for the resulting cloud software platforms market.

For private clouds, the inspiration came from server virtualization. Virtualization management frameworks were extended into private cloud frameworks and provided a path to evolve virtualization deployments into private cloud deployments. The underlying system software for these clouds consisted primarily of operating systems and hypervisors.

However, the influence of Amazon gave rise to a new class of system software that IDC calls cloud system software. While hypervisor management is typically constrained to the resources of single servers, cloud system software generally orchestrates resource management across multiple systems running open source or proprietary hypervisors. Most cloud system software solutions act as an overlay to existing virtualization and management infrastructure, replacing the user's management interface with one that provides a higher level of abstraction.

This gives the user the ability to orchestrate low-level resources into pools and often provides higher-level cloud APIs that other management software can use. Existing virtualization vendors and open source projects added these concepts to their platforms, and new vendors and open source projects designed using these principles have begun to emerge.

OPEN SOURCE AND ITS ROLE IN OPEN CLOUDS

Open source has long dominated public cloud platforms for many reasons:

- **Cost.** Many public clouds target consumers with free or low-cost services, and in the enterprise space, Amazon redefined the computing market with its utility-like model. Using open source software allows providers to create business models that can meet these economic requirements. The lower cost to operate is due to the provider's ability to self-support infrastructure and implement purpose-optimized solutions.

- **Customizability.** The software services of a cloud are the core and intellectual property of a cloud provider's business. Creating unique services at a reasonable cost often requires extensive customizability, and open source software enables this to happen.

- **Full visibility.** Having access to the source code allows cloud service providers to fully understand and debug the software if needed. Some providers that use open source never become heavily involved in modifying or developing the code, simply using vanilla builds of the software. However, many make use of access to the source code in order to self-support and customize the software because many cloud providers do not use a commercially supported version. Access to the source code allows providers to fully understand how the software works, tune performance, build custom software around it, and track down problems where they occur.

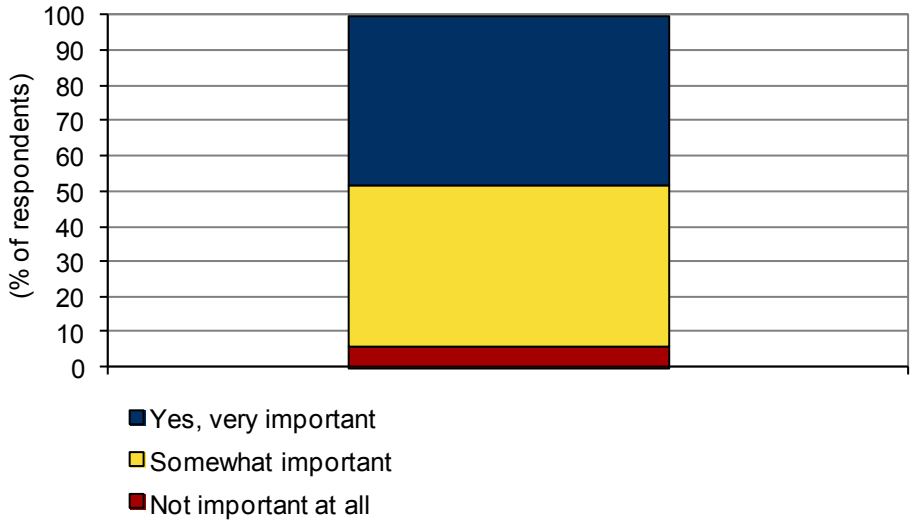
- **Collaborative development model.** The Linux and open source software communities have proven that they can build quality software and build it very quickly, and customers benefit from this virtuous cycle of development. Providers can be fully vested in the technology running their cloud systems and have direct access to the people building it. Open development in the cloud is one of the ways that the industry is addressing the increasing complexities of computing faster and more efficiently.

Figure 1 shows that over 94% of respondents felt that collaboration and a vibrant open source ecosystem were somewhat or very important to cloud adoption, with nearly 49% indicating that they were very important.

FIGURE 1

The Importance of Collaboration and Ecosystem

Q. Are collaboration and a vibrant open source ecosystem important for cloud adoption?



n = 282

Source: IDC's *2012 Cloud System Software Survey*

However, just because a cloud contains open source software doesn't mean that a cloud is open. It is not uncommon to find clouds that are compliant with the FOSS licenses of their components, but that have undocumented APIs that restrict either data portability or custom software integration for users. This is allowed under the terms of many licenses because service providers are not technically distributing the software, so whatever changes providers make are not required to be made available to users. In addition, configuration information (e.g., assembly, integration, and configuration of components) is not always public knowledge; therefore, someone else could not simply gather the same FOSS components and assemble them into a cloud. These considerations are important because the cloud service model also frequently emphasizes the ability to interact with cloud providers through standards-based APIs, and IDC's *2012 Cloud System Software Survey* revealed that openness is a high priority for a vast majority of respondents.

While public clouds have leveraged open source componentry to build their services, most enterprises will not do so for their private cloud for many of the same reasons they don't build their own Linux distributions. Even public cloud providers are beginning to look at open source cloud platforms in order to avoid reinventing the wheel time after time and focusing their development attention on areas they can differentiate. These customers will look to free or commercial prebuilt and consumable cloud platforms. Several open source cloud projects, such as CloudStack, Eucalyptus, and OpenStack, aim to provide these platforms for both public and private cloud customers.

Private cloud presents opportunities and challenges for open source. While enterprise virtualization is currently dominated by proprietary solutions, multihypervisor adoption is expected to double over the next couple of years as customers diversify and alternative solutions improve. IDC's *2011 Server Virtualization Multiclient Study* indicates that open source–based solutions are running neck and neck with proprietary solutions for customers' choice of a second hypervisor. In addition, not all private clouds will be built on top of existing hypervisor deployments. IDC's *2012 Cloud System Software Survey* indicates that while approximately half the market will use an existing hypervisor, the rest of the market will choose based on the cloud system software platforms and is willing to use whatever hypervisor is embedded. Therefore, while having a virtualization installed base will lead to private cloud opportunities, much of the private cloud market will also develop independently of virtualization. A large number of vendors, both start-ups and established companies, are seeking to take open source cloud system software into the enterprise with commercial offerings, imitating how Linux evolved from dot-com deployments and moved into the enterprise around the turn of the millennium.

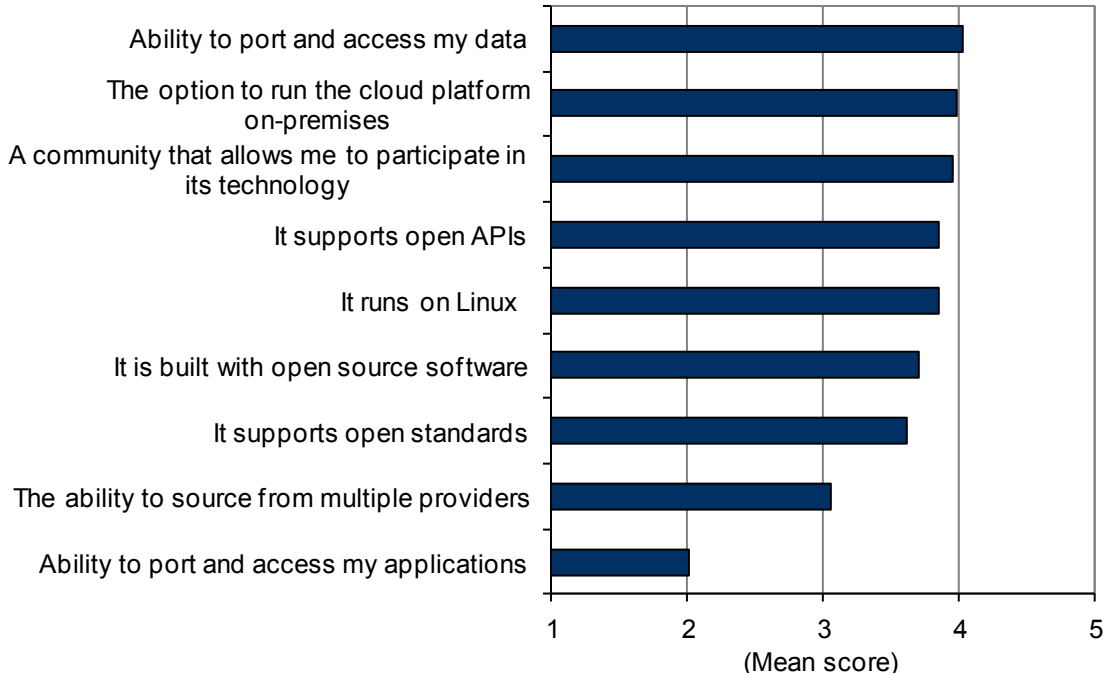
WHAT IS AN OPEN CLOUD?

Figure 2 shows how enterprise customers responded when we asked them to rate the importance of various characteristics that make an open cloud.

FIGURE 2

What Is an Open Cloud?

Q Please rate the following characteristics as to how important each one is to an open cloud, with 1 being not at all important and 5 being very important.



n = 282

Source: IDC's *2012 Cloud System Software Survey*

Enterprise Data Is Crucial

The results show that customers are most concerned with the portability and accessibility of their data, which is not surprising given that enterprise data is so unique and crucial to a business. Enterprises clearly also want a platform that they can consume on-premises so that they are not locked solely into public clouds. This fits in well with previous IDC data that shows most enterprises are investing more heavily in private cloud before venturing into public cloud.

Interestingly, the ability to port applications across clouds was rated very low by customers, with data being their primary concern. This may stem from the history of enterprise applications, which usually have very long lifetimes and, once they are built, stay on the platform on which they were built for the rest of their lives. It may be that many enterprises have just abandoned the idea of easily portable applications.

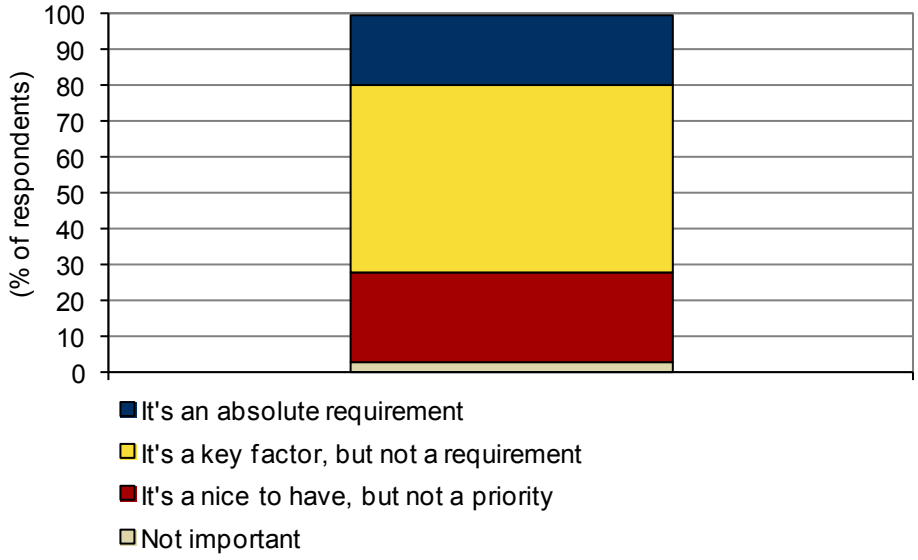
Open Source, Open Standards, and Open APIs

In terms of the software itself, customers rated open source software — particularly Linux — open standards, and open APIs important characteristics of an open cloud. Almost 72% of respondents indicated that when choosing a provider or building their own cloud, the use of open source software, open standards, or open APIs would be either an absolute requirement or a key factor (see Figure 3). Nearly 20% of respondents indicated that it would be an absolute requirement.

FIGURE 3

The Importance of Open Source, Open Standards, and Open APIs

Q. When considering a cloud provider or building your own cloud, how important is the use of open source software, open standards and/or open APIs to your organization/division?



n = 282

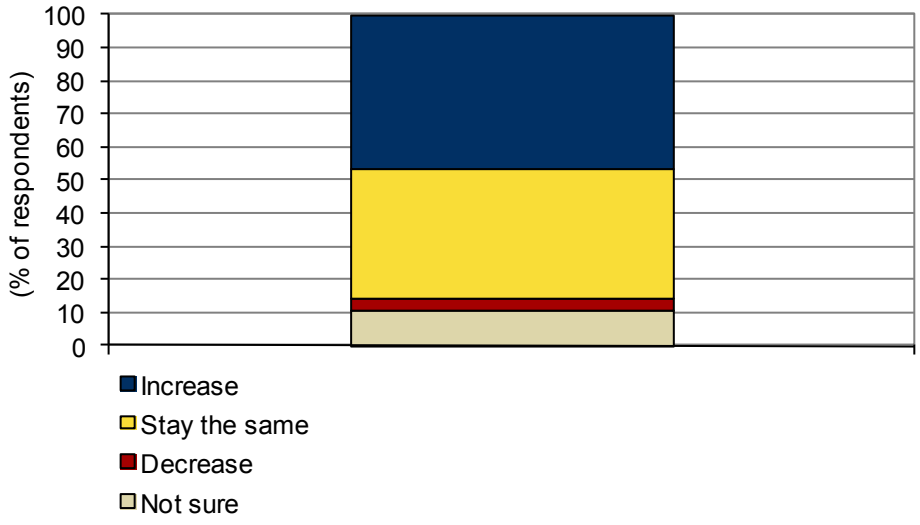
Source: IDC's *2012 Cloud System Software Survey*

Nearly 86% of respondents indicated that they would either increase or maintain their current spending on Linux and open source software in support of their private cloud plans, with nearly 47% indicating an increase (see Figure 4).

FIGURE 4

Spending on Linux and Open Source Software

Q. In the next 12 months, what best describes your plans for investing in Linux or open source software to support your cloud plans?



n = 282

Source: IDC's *2012 Cloud System Software Survey*

Community

One of the other highly rated open cloud characteristics was a collaborative community around the cloud. A strong community is a good sign of the sustainability and longevity of a project or service. Cloud is putting a higher emphasis on community for the following reasons:

- **Early maturity.** Private and hybrid cloud models are still stabilizing. An open collaborative model has been proven to work with Linux and was instrumental in ironing out early design and implementation issues. Cloud system software, similar to operating systems, will have diverse and broad deployment use cases and be put into a wide range of environments. A large range of issues, including hardware and software compatibility, will have to be addressed, and a diverse community of contributors will be needed to tackle the wide-ranging areas that cloud system software will touch. To replicate the self-support model that dramatically reduced the cost of using Linux, users will also need to connect with vendors, providers, and other users to share adoption and support best practices.
- **Broad functionality.** Cloud system software is on its way to becoming a universal software component and thus will provide a very broad range of functionality and serve a very broad range of use cases. Similar to how Linux has morphed into

many variants and diverse uses, so will cloud software. Being such a core platform requires that cloud software be developed with many requirements in mind and be compatible with a very wide range of hardware and related software. A strong community is needed for vendors and users to gather into these groups and take the software to where they need it.

A strong community can also help preserve the goals of an open cloud. Lock-in has been a pervasive issue since the dawn of IT, and cloud has the potential for greater abuse and lock-in than ever. Open source solutions can't be completely controlled by any one person or entity and allow each participant to have greater control over his or her own destiny. On the commercial side, they offer inherent transparency, multiple alternatives, and a self-policing community that ultimately serves the needs of everyone.

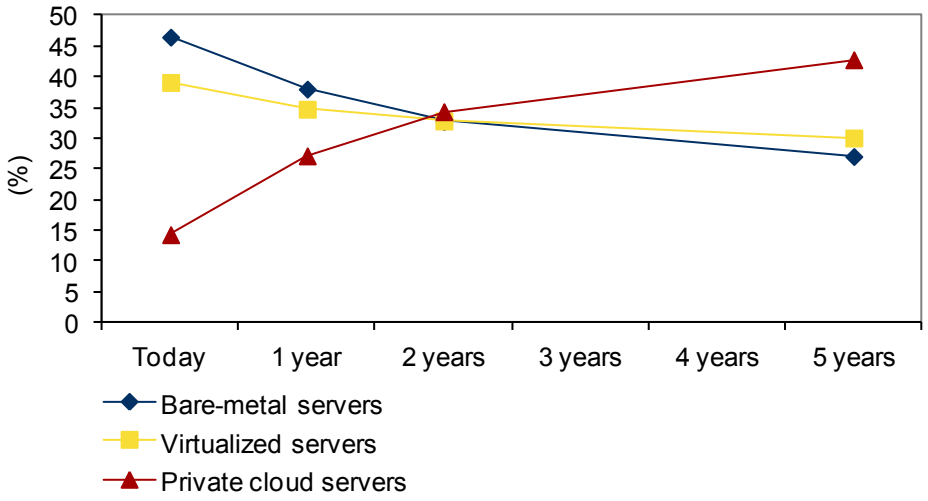
FUTURE OUTLOOK

Respondents feel that private cloud servers will constitute the majority of their infrastructure in five years, eclipsing bare-metal servers and traditional virtualized servers, as shown in Figure 5, leading to a large addressable market for cloud software platforms. While the exact timing can be hard for customers to predict, the results clearly show that the industry is moving to the private cloud model.

FIGURE 5

Private Cloud Takes over IT Infrastructure

Q. Please indicate the percentage splits of your current server infrastructure at present and going forward for the following categories.



n = 282

Source: IDC's *2012 Cloud System Software Survey*

Platform as a service (PaaS) and integration with public cloud generated a high amount of interest among respondents, with significantly more interest among current private cloud adopters, indicating that a cloud maturity model is forming where infrastructure as a service (IaaS) is adopted before PaaS and private cloud is adopted before public cloud.

CHALLENGES/OPPORTUNITIES

- **Existing versus new applications.** Respondents indicated that on average, 63% of the applications running on a private cloud will be existing applications. Many of these applications and their operating system or application frameworks will be closed source, and it will be important for open source cloud platforms to work with all environments, proprietary and open.
- **Hypervisor choice.** Today, enterprise datacenters are dominated by proprietary hypervisors, and nearly 47% of customers prefer to use their existing hypervisor. However, 53% of the market is open to adopting a new hypervisor that is either recommended for or integrated into a cloud platform. Thus, open source vendors can compete on the strength of their cloud system software and challenge the installed base of proprietary virtualization.
- **Self-service Web portals.** This feature of private clouds was rated significantly lower than other features in terms of satisfaction by current private cloud deployers. As the public face of a cloud deployment and with expectations set high by the likes of the Googles and Amazons of the world, this is clearly an area that vendors can use to differentiate themselves.
- **Shifting workload profiles.** As enterprises begin to refine their allocations of workloads between public and private, some traditional areas of strength for Linux and open source, such as Web infrastructure, are being moved into the public cloud. Respondents indicated that business processing (ERP, CRM, OLTP, batch) and decision support (data analysis/mining/warehousing) would be some of the top workloads for private cloud systems, requiring the open source market to adapt to a new workload profile.

CONCLUSION

Open source cloud software has been tremendously successful in public clouds and is now poised to penetrate the enterprise private cloud market. While some private cloud solutions will evolve from existing virtualization deployments, an equally large part of the market will be built "greenfield" using cloud system software and embedded virtualization technology. Open source software, Linux, open standards, and open APIs are all strongly viewed as being crucial characteristics of an open cloud. Results from the survey indicate that for open cloud to be successful, a strong community and ecosystem is one of the top factors in maintaining openness and increasing adoption.