I D C   E X E C U T I V E   B R I E F

Optimizing Information Management in the Cloud

June 2011

Introduction

Firms are going through an evolutionary transformation to a cloud-based delivery model. Both technology and business factors are influencing this transformation. The advent and proliferation of virtual infrastructure, which separates logical from physical structures, is the cornerstone of this transformation. Other technology trends contributing to this transformation include IT service automation, converged infrastructure, scale-out server and storage architectures, and unified management of different technology domains. These are all enabling components of a cloud-based architecture.

However, technology is only part of the equation to this transformation. The economics of both IT and storage infrastructure are driving change. Firms seek to curb the percentage of the overall IT budget that is consumed by capex and opex investments in storage. Firms must transform their datacenter operations from an infrastructure focus to a service-level focus, particularly in light of how the cloud impacts information management. Firms must address the operational imperative of ensuring that the right service levels are provided to a business unit or application. The advancement in new technology and the integration of this technology make running IT as a service, either in a public or private cloud setting, increasingly more viable.

This Executive Brief describes the impact that both virtualization and the cloud are having on organizations and discusses the subsequent effects on information management.

Leveraging Virtualization and Cloud to Transform Business

2010 was a pivotal year in the use of server virtualization technology. According to IDC, it was the first year in which more new application instances were deployed as virtual machines on a virtualized server than were deployed on a dedicated physical server. Based upon typical server replacement cycles and new application deployments, IDC believes that the majority of workloads will be soon be running as virtual machines (see Figure 1).

Virtualization Is the Foundation Platform for the Datacenter

Worldwide Installed Workloads Virtualized by Year
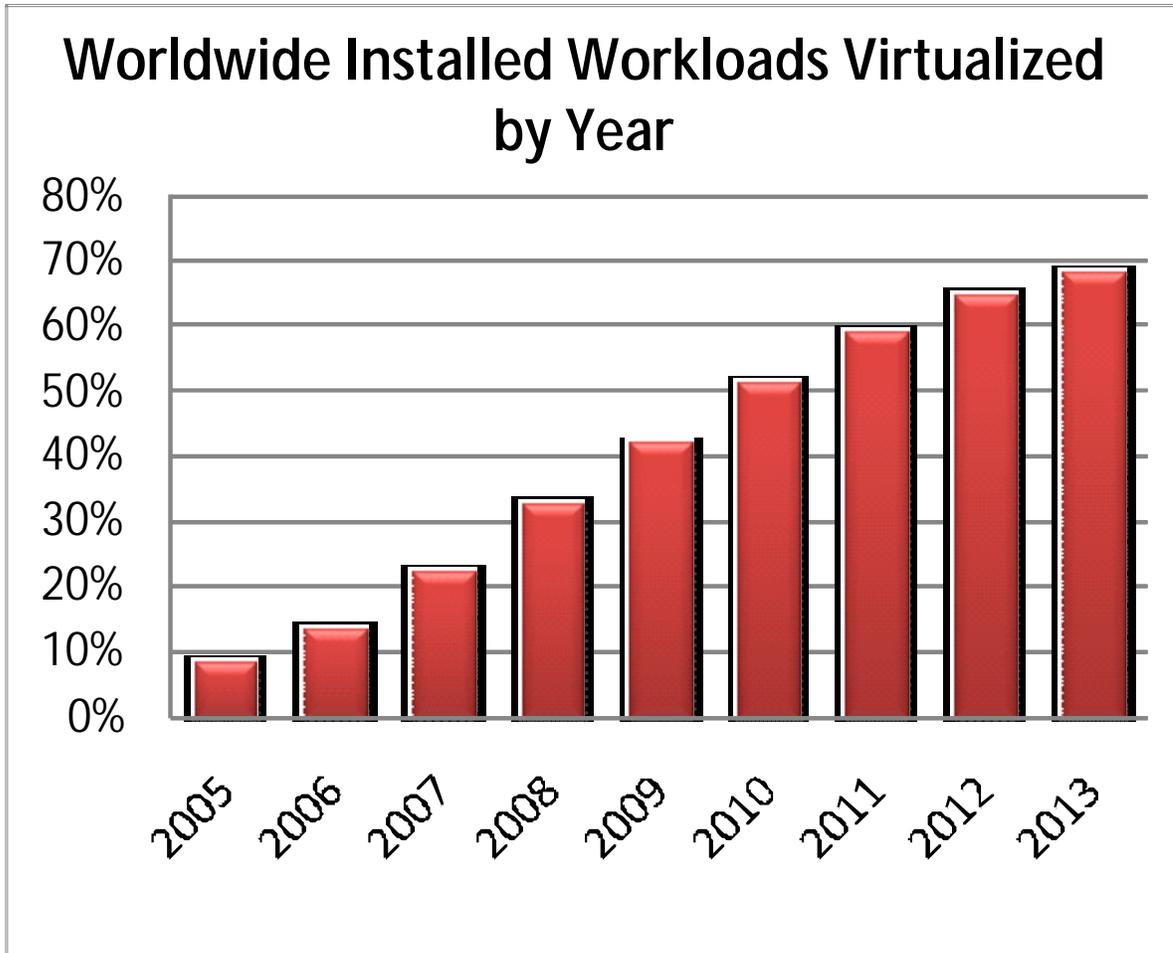
Source: IDC, 2011

Advancements being made in the datacenter with respect to consolidation, virtualization, and improvements in management, utilization, and efficiency are now spurring interest in "private" cloud environments designed to provide a more holistic and dynamic IT environment that can adapt to changing business needs. At the same time, organizations are also leveraging public cloud offerings that are an extension of and complement to the internal datacenter. Organizations are pursuing these strategies because they are seeking to address unpredictable data growth, operational pressures (e.g., offloading certain provisioning, reporting, or administrative processes), and business pressures (e.g., faster exploitation of new opportunities and quicker response to customer demands).

New Realities of Virtualization and Cloud Shape IT Requirements

A variety of factors are combining to influence how virtualization and cloud are deployed and managed by organizations in today's fast-paced, information-centric environments. CIOs seek to improve the availability of business-critical applications and the ability to quickly and easily meet

regulatory mandates. Business units demand ubiquitous and constant access to information (including on new mobile devices) from their IT departments.

The IT teams tasked with delivering on these business and technology requirements must respond to senior management demands for improved asset utilization, the unprecedented growth in unstructured data, concerns about data security, and budget constraints that affect investments in capital and staff. IT departments need to lower IT costs and ensure the backup, recovery, archiving, and discovery of information in the datacenter or in the cloud.

All of these demands are placing significant strain on IT organizations that must manage the compute, storage, and networking infrastructure to deliver the applications and data necessary. It is no surprise that IT organizations are increasingly turning to virtualization and the cloud to address operations and information challenges. Among the specific goals that IT seeks to achieve with virtualization and the cloud are the following:

- Evaluating advanced cloud (private and public) services to further reduce costs and boost responsiveness

- Aggressively adopting cloud-based archiving solutions to better address the retention and discovery of content, including emails, files, Web 2.0 content, medical records, and so on

In this fast-evolving environment with a mix of legacy physical systems, next-generation virtual assets, and use of cloud-based services, a major concern for IT organizations is whether they can confidently take advantage of these new solutions without jeopardizing data protection and information security standards. Among the specific issues IT needs to address are the following:

- Can IT reliably protect and archive physical and virtual machines using well-tested backup solutions and practices?

- Will the use of cloud-based applications exacerbate or reduce the operational impact and the complexity of monitoring, alerting, and reporting on information and infrastructure assets across physical, virtual, and cloud platforms?

Greater Expectations and New Challenges for IT Organizations

For traditional environments within the datacenter, IT organizations made significant investments in money and staff training to reach optimum levels of control for individual, stove-piped IT assets. The transformation to dynamic IT and the use of cloud delivery models pose risks to maintaining this control. The modern IT environment, with its greater use of virtualization and cloud-based storage resources, profoundly affects the way IT is managed and delivered. From a technological and operational perspective, IT organizations face increasing expectations when it comes to:

- Providing end-to-end data protection and backup/recovery services that support new models (virtual and cloud) while meeting more stringent application and user recovery requirements

- Obtaining control over the incessant growth of information with respect to long-term management, security, and value

- Extending access of applications and data to expanding numbers of internal and external users, with mobility a growing and pervasive security and control challenge

- Improving asset utilization across the increasingly complex set of physical and virtual assets as well as internal and cloud environments

As IT considers the broader use of virtualization (private cloud) and public cloud–based services, IT managers need to ensure optimum levels of control with respect to performance, reliability, efficiency, and security among applications and end users. These challenges can be categorized into three broad segments:

■ Loss of control

- Security and control of data — both in motion and at rest (e.g., Where does the data reside, and how is security tested?)

- Uncertain recovery across physical and virtual systems at multiple locations (e.g., Have local teams tested recovery times and points in new environments?)

■ Reduced reliability

- Performance — unpredictable application performance and backup times

- Lack of integration with existing application and infrastructure recovery processes

■ Increased uncertainty

- Complexity of cloud choices, service provider lock-in, and unpredictable pricing

- Inability to integrate new cloud-based services with existing applications and processes

Gaining Control with Better Information Management

Any effort to speed the adoption of virtual servers and use of cloud-based IT services depends upon IT organizations' confidence in their ability to lower IT costs and confidently back up, recover, archive, and discover information anywhere. This requires the use of a flexible and scalable information management solution that automates data protection, recovery, and archive processes while enabling policy-based management across both internal and external (cloud) IT assets. Key processes that information management solutions must address include:

■ **Backup and recovery.** Securely moving business data and virtual machine images between on-premises systems and hosted or cloud-based facilities to reduce remote site administration costs and long-term tertiary storage costs

■ **Archiving.** Long-term, low-cost retention and/or preservation of information through policy-based data movement and data deduplication services that link existing internal applications with cloud-based application and archival services

■ **Data retrieval and ediscovery.** Efficient and timely retrieval of locally and remotely archived content, including emails, files, Web 2.0 content, and medical records for business or litigation purposes

The most important criterion is that this information management solution deliver these capabilities for both virtualized and cloud deployments, no matter which approach IT chooses to take today or in the future.

IDC believes that data deduplication dramatically improves the economics of a customer's overall virtual and cloud infrastructure by minimizing storage footprint requirements, backup windows, and network bandwidth consumption. By reducing the data footprint, data deduplication also alleviates the need for additional spend on staff and storage hardware while preserving valuable server CPU cycles and improving performance, especially in distributed environments.

IT managers need to ask themselves if the solution they are choosing allows them to expand the use of data deduplication for backup and retention. Does it ensure security of data in motion and at rest? Does it improve archiving organization and retrieval processes? Does it enable quick incorporation of new

cloud applications into the IT environment? IT organizations will need a solution supplier that provides an integrated set of management offerings (both software and cloud based) that allow IT to monitor and report on the overall information environment — both inside the datacenter and in the cloud.

Conclusion

Large-scale virtualization as well as greater use of both private and public cloud services will deliver many benefits to organizations in terms of cost reductions and operational efficiencies. To reap all these benefits, however, organizations must assess the impact such technologies have on current and future information management processes such as backup/recovery, data retention/archiving, and information security/discovery.

IT managers need to actively evaluate vendors in this market and look for a partner that delivers robust and flexible data deduplication solutions while allowing IT to achieve all the benefits of virtualization. This information management partner must be able to improve the efficiency of backup, archiving, and ediscovery efforts across a wide and growing variety of virtual and cloud deployment/use models.