WHITE PAPER

The Importance of Data Control in Hybrid IT

Sponsored by: NetApp

Ashish Nadkarni
Laura DuBois
Richard L. Villars
July 2014

EXECUTIVE SUMMARY

As businesses become data-driven entities, they are holding their CIOs (and in turn information technology [IT] organizations) accountable for optimizing business success. IT organizations are rapidly moving from a static to a more dynamic, services-based model in response to businesses' growing demand for faster creation and delivery of services to end customers. One consequence of this move is the emergence of hybrid IT organizations that can improve datacenter asset utilization in their own facilities while also leveraging cloud-based assets to extend services and capabilities.

One of the ways in which CIOs can be successful in supporting the goals of their business is by controlling data. The management of data growth while also enabling the business to gain more insight from the data itself is the key challenge for a hybrid IT organization. IDC believes that controlling data flow in a hybrid IT environment is crucial for businesses to effectively transform themselves into data-driven entities. According to IDC, *Data Control* means addressing the following:

- **Business needs:** Enabling data access and intelligent data placement across multiple internal and external data assets to mitigate latency and accelerate use of data analytics in business processes
- **Information technology needs:** Enabling dynamic and automated deployment and management of infrastructure based on costs, performance, scalability, and data security
- **External needs/mandates:** Protecting data from unauthorized outside access or manipulation and managing data assets to meet governance, privacy, and regulatory requirements

Data Control was relatively easy when all the data resided inside the internal datacenter – which was often owned and managed by the business. As CIOs move important business assets to third party-owned collocation, hosting, and cloud facilities, Data Control becomes increasingly challenging. The only way that CIOs – as "Chief Data Stewards" – can manage all three facets of Data Control in this new "cloud enabled" era is by extending Data Control across all the assets on which corporate data resides. Data Control has to be dynamically adaptable and agile and offer data portability across local datacenters and the cloud. In other words, Data Control should work in a hybrid IT environment. Very few suppliers can lay claim to the ability to offer consistent and universal Data Control for hybrid IT environments. NetApp offers an ONTAP-based Universal Data Platform that provides dynamic data portability and extended customer choice for storage in hybrid IT deployments. More importantly, NetApp's solution is about Data Control without compromises in choice.

THE NEW DATA-DRIVEN BUSINESS ERA

Businesses everywhere are in the process of becoming data driven. While the term *data driven* is a bit amorphous, in the context of Data Control, it basically means intelligently controlling every business decision based on data patterns and analytics. We live in a connected world in which businesses have no choice but to quickly adapt:

- **Mobile:** The pervasiveness of mobile devices and smartphones the world over means that data is generated from a lot of end devices. More users mean more data sources.
- **Social:** Users are leveraging social media to engage with each other, and businesses want in on that interaction. They want to learn from user-to-user interaction to offer better products and services and/or use the data to sharpen their marketing or sales efforts.
- **Internet of Things:** Sensors and machines now have capabilities to report data, which businesses can analyze and proactively act on. Or in many cases, they can use this data to improve service quality or operational efficiencies.

In short, businesses want to collect more data from more sources and analyze all of it to change user experiences and business outcomes. In many cases, the change rate of this data is very high or the shelf life of this data is very short — which means businesses have to analyze this data quickly before it changes or gets too old. They also want to store more data and store it for a lot longer in the event that it is needed for further analysis in the future.

The Era of Hybrid IT

IT organizations in data-driven businesses are rapidly moving from a static delivery model to a more dynamic, services-based model. In the static delivery model, IT took a reactionary approach to the delivery of information technology. This model is cumbersome and archaic and does not fit well into businesses that heavily lean on data-driven mobile, social, and analytic techniques to stay competitive.

A dynamic, services-based approach allows IT organizations to effectively service the "data needs" of their internal clients — business units to be specific — in the same way the business as a whole services its external clients.

This dynamic, services-based approach is known as hybrid IT and uses cloud-like (i.e., as-a-service) delivery as the overarching framework, but it leverages both on-premises datacenter and public cloud assets. It decouples data delivery from the physical assets upon which this data resides. IT organizations can then leverage both internal (their own datacenter-based) and external (public cloud-based) assets to achieve a hybrid IT model that delivers:

- **Autonomy:** Data delivery assets are autonomously provisioned and/or managed. Instead of an application dependency on the infrastructure, there is now a framework that seamlessly bridges disparate assets into a single data delivery entity.
- **Agility:** Performance and capacity requirements are proactively accounted for. Changes to the data delivery framework are dynamically and nondisruptively adjusted. Internal clients don't need to intervene operationally.
- **Innovation:** By decoupling the physical assets from data access, IT organizations can insert innovation into both their delivery framework and the infrastructure.

The immediate impact of hybrid IT is that traditional enterprise datacenters are evolving into private cloud datacenters. They are also being virtually expanded into public cloud. In other words, applications and their associated data sets can bidirectionally flow between the private cloud datacenter and public cloud assets.

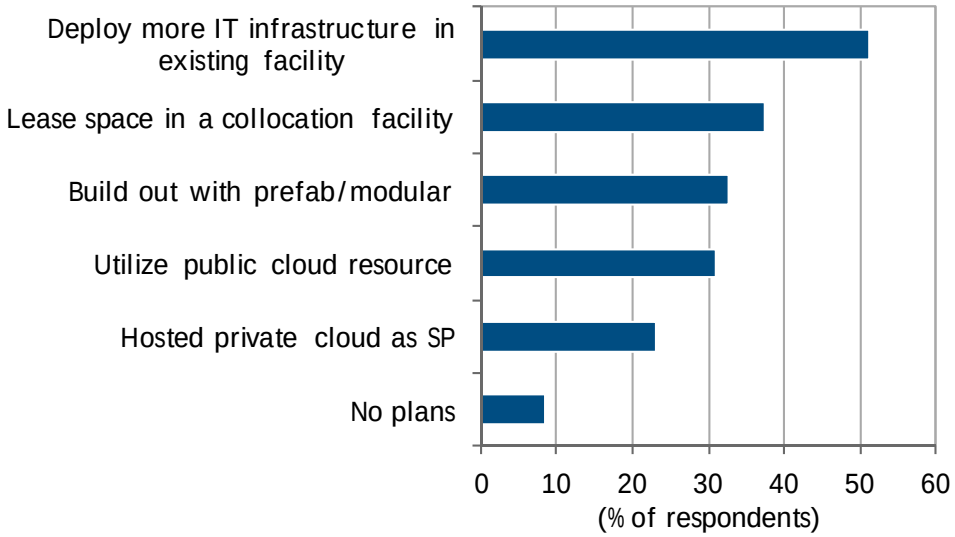
In a hybrid IT world, CIOs must instill a service provider culture in their organizations. Furthermore, as stewards of their businesses' data, CIOs need to switch their operational mind frame to "Data Control in Hybrid IT."

What IT Executives Are Saying About Datacenter Expansion

In a survey conducted by IDC in December 2013, a majority of United States-based IT executives stated that they are planning to extend their datacenter capacity by moving assets to a collocation facility, leveraging a hosted private cloud, or using public cloud services (see Figure 1).

FIGURE 1

Plans for Adding Additional Datacenter Capacity over the Next 18 Months



n = 410 U.S. IT executives

Note: Multiple responses were allowed.

Source: IDC, 2014

Data Control in Hybrid IT

For CIOs of IT organizations (or those seeking to become CIOs) the role of Chief Data Steward means that Data Control now becomes a core focus area. For most businesses, Data Control means satisfying the following needs:

- **Business needs:** Enabling data access and intelligent data placement across multiple internal and external data assets to accelerate use of data analytics in business processes
- **Information technology needs:** Enabling dynamic and automated deployment and management of infrastructure based on costs, performance, scalability, and data security
- **External needs/mandates:** Protecting data from unauthorized outside access or manipulation and managing data assets to meet governance, privacy, and regulatory requirements

Data Control starts with taking a data-centric approach to asset management. By empowering IT organizations to optimally manage all facets of Data Control, CIOs can seek to support business needs and therefore optimize business success.

Data Control should be looked at not as a singular entity but as a framework. Indeed, in larger enterprises with several hundred applications or more, Data Control can be a challenging proposition with varying types and levels of control based on the mission or business criticality of the applications in question. Effective Data Control therefore involves effective tiering and management of data needs across the entire business application spectrum.

For an IT organization that is seeking to morph itself into hybrid IT, direct and indirect costs of losing the Data Control framework can be prohibitive. Incurring such costs can be as bad as not implementing Data Control at all. Losing Data Control is akin to an aircraft losing control midflight. Risks of losing control include:

- **Higher IT costs:** Every dollar that is spent by IT impacts the business' bottom line. Lack of or no Data Control leads to higher operational overhead and therefore higher people and infrastructure costs.
- **Reduced insight into data:** Lack of effective data management leads to reduced insight into the type of data generated by the business. This can lead to inefficient data placement, exposure to data loss, or even situations where data is not disposed of in an appropriate manner, leading to more infrastructure costs.
- **Data governance and compliance risk:** Businesses that are subject to governance, risk, and compliance (GRC)-type oversight can pay a heavy price for not maintaining data in an appropriate manner.
- **Business impact:** Businesses can lose their competitive advantages because of the lack of or absence of Data Control. Loss of competitive advantage can be caused by faster time to market for new products or services (compromising quality), data-driven decisions that are poor or ineffective, and ultimately loss of customer trust.

Challenges with Data Control in Hybrid IT

Datacenter-based and cloud-based environments are the foundation of hybrid IT — primarily because they allow IT to offer an agile and dynamic infrastructure to satisfy the appetite of data-driven businesses.

Hybrid IT environments are made of hosted private and public cloud-based systems in addition to on-premises private cloud and traditional IT environments. Cloud-based systems alter the control dynamics of the infrastructure. They introduce a situation where IT has to effectively control data regardless of where the IT assets are located (in an internal datacenter or a third-party cloud datacenter). In other words, in hybrid IT, Data Control cannot be outsourced.

This situation is compounded by the fact that various business units are clamoring for increased access to and visibility of their data. With this access, they can perform time-sensitive analytics — which in turn results in more data being generated. In all of this, IT has to ensure that the data is stored in a compliant and secure fashion.

Cloud-based infrastructure can pose an extra set of challenges to Data Control. These include:

- **Managing disparate storage services:** IT organizations have to deal with different types and classes of storage services that can differ widely from provider to provider and even within their own on-premises storage infrastructure.
- **Data mobility and visibility:** Moving data between different service tiers located either on-premises or in the cloud can be challenging. Data visibility can be crippled if automation is introduced without effective data tracking mechanisms.
- **Limited choice with service providers and/or the perception of provider lock-in:** Several businesses consider public cloud services to be a form of outsourcing, increasing the perception that once data is moved into the cloud, it stays there forever. The fact that the number of providers offering comprehensive storage or data services remains limited reduces the options businesses have with the cloud.

NETAPP ENABLES DATA CONTROL IN HYBRID IT

The key to addressing the challenges discussed previously is the development and deployment of Data Control — a proven way to effectively manage and maintain data in hybrid IT environments. Not many suppliers can claim to offer a solution for implementing consistent Data Control across all of IT. This is because such a deployment requires the use of a single storage management platform that can be placed inside a datacenter and in the cloud to manage data on-premises and in the cloud. In other words, the platform itself has to be universal.

NetApp is one of the few suppliers that can offer a solution to implement dynamic Data Control in hybrid IT environments. NetApp's Universal Data Platform is designed as a cloud-ready platform. It offers features such as:

- **Standard data containers:** Consistent service quality across all cloud and datacenter assets, including public cloud services
- **Secure multitenancy:** Ability to host multiple workloads belonging to multiple business units without any service degradation or data boundary compromises
- **Pooled virtual resources:** Ability to abstract hardware resources irrespective of their location or type
- **Efficient data transport:** Efficient local and wide area data mobility transport
- **Rich data management features:** Feature-rich data management capabilities

NetApp's Universal Data Platform also works across NetApp FAS arrays as well as third-party arrays and even commodity disk hardware. It can be placed in the datacenter, next to the cloud, and in the cloud. As a common storage OS powering NetApp's FAS platform, the Universal Data Platform offers a level of pervasiveness not offered by many suppliers. This allows IT to provide standardized data services across the entire hybrid IT infrastructure. These include data protection, provisioning, efficiency, replication, and quality of service. In addition, the NetApp Universal Data Platform offers:

- **Dynamic data portability — the ability to efficiently move data between clouds using time-tested native replication capabilities.** The fact that it offers multiprotocol data connectivity means flexibility in connecting multiple environments on-premises and in the cloud.
- **Extended customer choice — the ability to choose from several public cloud service providers without worrying about specific offerings or provider lock-in.** The Universal Data Platform supports all major virtualization and cloud management platforms, allowing IT organizations the flexibility in choosing where, how, and with whom to deploy.
- **Seamless connectivity among clouds — the ability to implement a multicloud strategy.** The Universal Data Platform offers features such as nondisruptive operations (NDOs) along with industry-leading scalability and efficiency capabilities, allowing businesses to rightsize their IT infrastructure. More importantly, they can now place applications and their associated data sets on the appropriate "cloud tier" that is governed by the Universal Data Platform.

Above all, one of the key differentiators of the NetApp solution is Data Control without compromising choice. Several suppliers are taking a captive approach to addressing hybrid IT. In other words, their solutions do not offer platform flexibility when it comes to deploying Data Control. NetApp, on the other hand, is focusing on helping CIOs maintain Data Control on their terms — by allowing them to choose the public cloud service providers they want to work with and only when they need their services.

Northumberland County, Pennsylvania, Embraces Hybrid IT

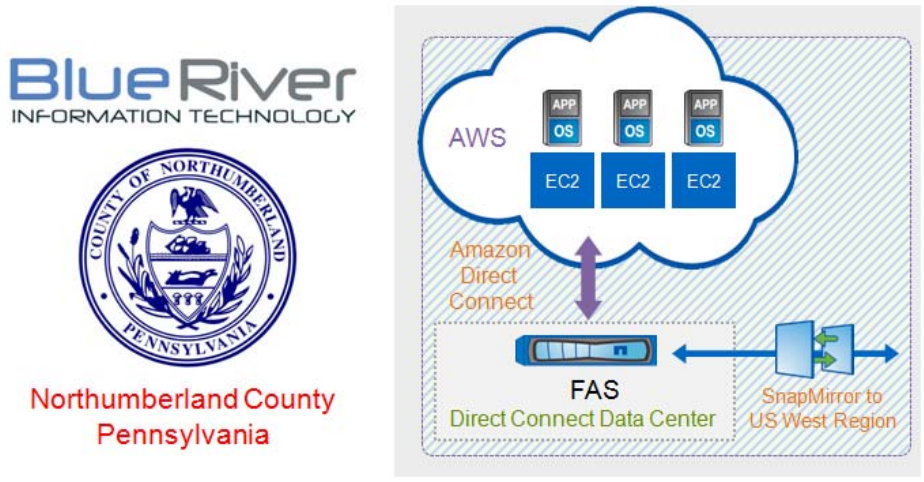
Northumberland County is located in the state of Pennsylvania. As of the 2010 census, it had a population of 94,528. The IT organization for the county faced typical problems shared by many state and local government IT organizations:

- Its datacenter was aging and in dire need of infrastructure upgrades. The datacenter facilities were also at risk for flooding.
- It had legacy IT systems that limited workforce productivity and hence reduced the effectiveness of county services. Frustrated by lack of functional IT systems, much of the workforce resorted to paper-based systems.
- It had severe budget limitations, meaning that it could not invest in datacenter infrastructure or system upgrades, leading to further deterioration of the situation.
- It faced regulatory requirements that meant adhering to data retention and protection policies.

Frustrated by the situation, the county IT-in-charge sought help from Blue River IT, a local reseller. Blue River IT designed a hybrid IT infrastructure that leveraged a clustered ONTAP solution with Amazon Web Services (AWS) as the public cloud solution. By moving to a collocation facility, the county was able to leverage the benefits and economics of AWS cloud services while maintaining Data Control by connecting on-premises clustered Data ONTAP FAS arrays to AWS EC2 instances. It also leveraged SnapMirror to replicate the FAS arrays to a datacenter on the West Coast for disaster recovery (see Figure 2).

FIGURE 2

Northumberland County Hybrid IT Architecture



Source: NetApp, 2014

Nanyang Technological University Adopts Hybrid Cloud for HPCC

Nanyang Technological University (NTU) is one of the two largest public universities in Singapore. Over the years, NTU has grown to become a full-fledged research university, with a student population of around 33,000, focusing on research in the areas of energy, water, life sciences, material sciences, computational chemistry, and defense.

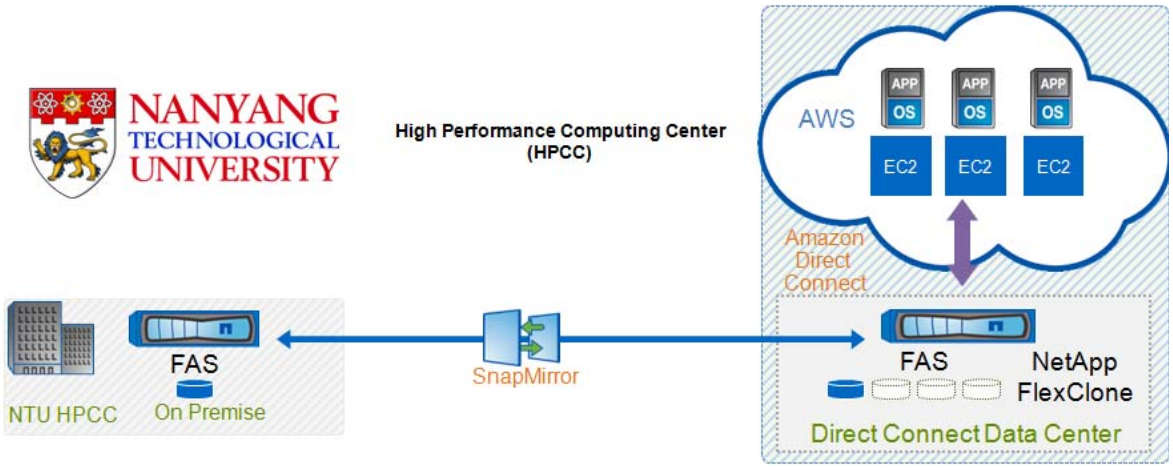
NTU's High Performance Computing Center (HPCC) implemented a Red Hat OpenStack-NetApp hybrid cloud solution (see Figure 3). This solution included:

- NPS for AWS hybrid cloud solution for cloud bursting and disaster recovery
- Red Hat CloudForms for seamless management

NTU HPCC leveraged this solution for HPC workloads such as Typhoon-Sim. The solution resulted in the center saving nearly 30% in operational costs while gaining greater Data Control. The center received a Red Hat Innovation Award in 2014, the first such win for an ASEAN university.

FIGURE 3

Nanyang University HPCC Architecture



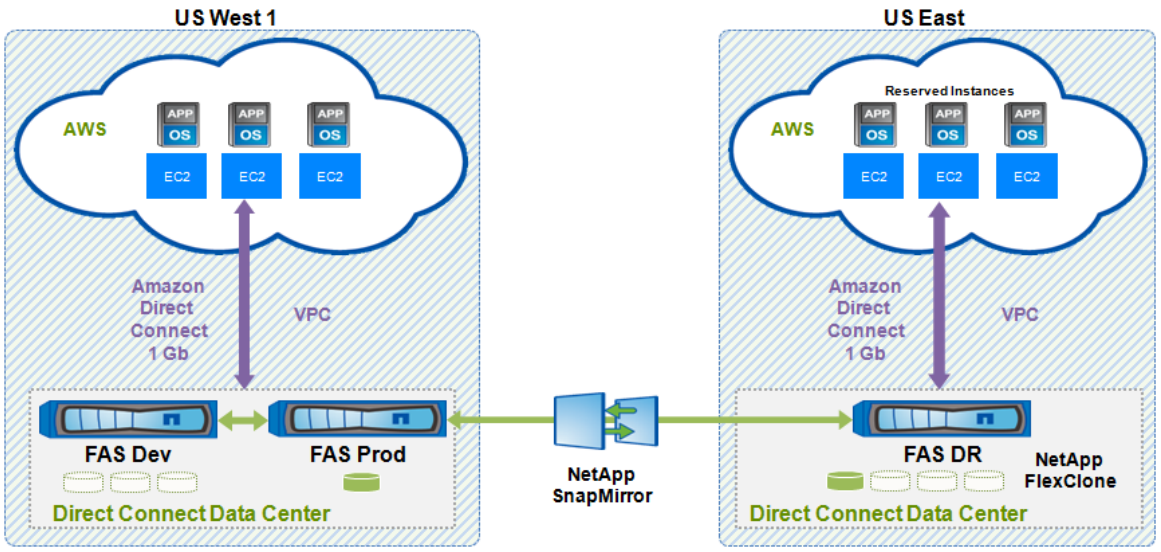
Source: NetApp, 2014

Leading Financial Services Firm Adopts Hybrid IT for Payroll Application

When a leading financial services firm was exploring systems architecture options for its payroll application, it chose a hybrid cloud approach. With this approach, the firm was able to migrate 90% of the virtual machines (VMs) to AWS EC2 instances, thereby saving 70% of server costs. In the process, the firm managed to address disaster recovery requirements in a cost-effective manner while staying compliant. Using this hybrid IT approach, the firm has achieved its goal of effectively managing nearly 10,000 virtual server instances with "90% of the benefits of public cloud with 10% risk" (see Figure 4).

FIGURE 4

Leading Financial Corporation Payroll Application Architecture



Source: NetApp, 2014

OPPORTUNITIES AND CHALLENGES FOR NETAPP

The IT landscape is rapidly evolving. Businesses everywhere are embracing or planning to embrace the public cloud, but for various reasons, many of them are not yet ready to give up their own on-premises infrastructure. Such businesses will no doubt appreciate the value of Data Control and the benefits of implementing Data Control across their infrastructure. NetApp needs to aggressively pursue:

- **Businesses that have already embraced a hybrid IT model.** NetApp needs to create a hybrid IT environment to more quickly identify where these businesses can attain immediate benefits from implementing Data Control with the Universal Data Platform.
- **Businesses that are still evaluating the option of adding public cloud to their infrastructure.** NetApp needs to articulate how a Universal Data Platform implementation can address the data governance and control concerns that are preventing these businesses from exploiting cloud services.

NetApp needs to aggressively preach the benefits of a storage platform that works across multiple on-premises and cloud infrastructures by leveraging its brand and singular storage OS mindshare to further the concept of Data Control. Most importantly, it needs to be selective with its portfolio of public cloud partners, showcasing successful Universal Data Platform–based hybrid IT deployments and providing IT organizations with more options.

ESSENTIAL GUIDANCE

Public cloud services are here to stay, and for most businesses, they offer an attractive value proposition. For many businesses, however, giving up investment in their own datacenter infrastructure is not yet an option. By hurriedly extending their infrastructure to the public cloud without implementing an overarching data management domain, IT organizations can quickly expose their business to increased risks. The resulting mess when it comes to Data Control can be worse than not adopting the public cloud at all.

CIOs should precede embracing of public cloud services by properly planning and then implementing Data Control in their IT and private cloud environments. Data Control must factor in business needs, information technology needs, and any external needs such as governance, regulatory, and/or compliance requirements. With the selection of a supplier such as NetApp that has partnerships with a growing set of public cloud providers, a CIO can use Data Control to better take advantage of rapidly evolving hybrid IT.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com