Today's enterprises need to be agile to increase the speed of innovation, improve responsiveness to customer demands and keep up with the continually changing business landscape. Along the journey to improve organizational agility, enterprises are increasingly embracing microservices-based, cloud-native applications, leveraging both on-premises data centers and public cloud resources to meet diverse geographic, data sovereignty and application dependency requirements.

However, running a business in a hybrid cloud environment is challenging as it introduces the complexity of managing containerized applications and data spread across multiple locations and types of infrastructure. Many cloud providers offer their own data services, but they are only available on their own cloud platform and not across hybrid cloud environments.

According to the following report from 451 Research, part of S&P Global Market Intelligence, "An increase in data processing location choices requires enterprises to make tactical and strategic decisions about when and where to process data, as well as to formulate strategies for data movement or federated data processing and analysis." For this reason, enterprises need to adopt a cloud native data plane solution that can seamlessly manage and move applications and data across different on-premises and cloud environments.

This report highlights the need and value of an infrastructure-agnostic strategic data platform to deliver on the promise of a true hybrid cloud with full application flexibility.
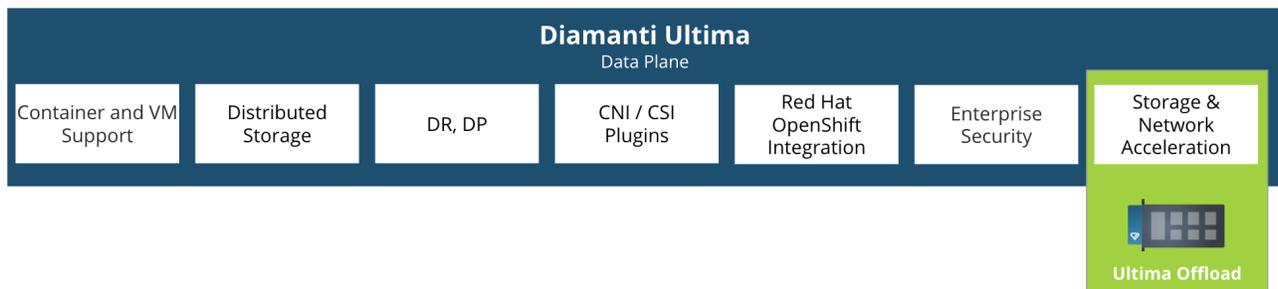
Diamanti is solving the challenge of container-based hybrid clouds with a simple and cost-effective platform that addresses enterprise requirements for managing Kubernetes data at any scale.

Simplify Data Management in Enterprise Kubernetes

Diamanti Ultima is the only data plane software with integrated networking and storage that delivers cloud native data services that extend across on-premises and cloud-based environments, supporting any Kubernetes cluster distribution. It features advanced Container Network Interface (CNI) and Container Storage Interface (CSI) plugins to deliver the following:

- A common operating model for managing data services across different clouds and infrastructure

- Container-aware data services for the protection and resiliency of applications, including:
    - Instant zero-copy volume snapshots
    - Off-cluster backup and restore
    - Asynchronous volume replication for fast cross-cluster disaster recovery and availability
    - Synchronous volume mirroring within clusters, even those spanning across availability zones

- Hybrid cloud data portability empowering enterprises to seamlessly migrate and failover stateful applications between clusters spanning across hybrid clouds while maintaining persistent data

- Interoperability with various public clouds as well as virtual and bare-metal infrastructures

- Support for various Kubernetes management solutions including Diamanti Spektra, Red Hat OpenShift, Amazon EKS and other Kubernetes platforms

- Boosts the performance of on-premises infrastructure by offering an offload card to accelerate networking, storage, management and security

Figure: Diamanti Ultima at a glance

Diamanti Ultima Data Plane						
Container and VM Support	Distributed Storage	DR, DP	CNI / CSI Plugins	Red Hat OpenShift Integration	Enterprise Security	Storage & Network Acceleration
						Ultima Offload

For more information about Diamanti and its products, please visit diamanti.com.

The case for an infrastructure-agnostic hybrid IT and multicloud data platform layer

JANUARY 21 2021

By Matt Aslett

The database services offered by the cloud giants are becoming increasingly attractive as more data is stored in the cloud, but there are also arguments in favor of an infrastructure-agnostic data platform layer that can run on multiple clouds, as well as in on-premises datacenters.

451 Research

S&P Global
Market Intelligence

Introduction

Enterprises today face a bewildering array of data platform options, including numerous data platform specialists, as well as a variety of services from the major cloud providers. As an increasing proportion of data is stored and processed in the cloud, the database services offered by the cloud giants become increasingly attractive, but there are also solid arguments in favor of adopting an infrastructure-agnostic data platform layer that can potentially run on multiple clouds, as well as on-premises datacenters.

The choice is not a zero-sum game: Many enterprises have multiple data platform providers and will be utilizing specific data platform products and services, including those from multiple cloud providers, for multiple individual use cases. However, when it comes to strategic choices, there are multiple arguments in favor of adopting an infrastructure-agnostic strategic data platform provider that is able to address hybrid IT and multicloud availability.

451 TAKE

Support for hybrid IT and multicloud is not required for all database workloads. As such, there are plenty of good arguments for taking advantage of data platform services that are only available on a single cloud platform, either tactically for a specific use case or strategically if a customer has made a significant commitment to a specific cloud platform provider. However, our survey data shows that nearly three-quarters of enterprises use multiple cloud providers, and a similar proportion are deliberately adopting a hybrid IT strategy to take advantage of both on-premises and cloud computing resources. While not all components in a software stack need to be available across multiple computing platforms, there is a case to be made that data processing is strategically important enough to justify investment in an infrastructure-agnostic data layer.

Context

451 Research survey data indicates that, for most companies, the future of IT architecture is a hybrid of on-premises and cloud resources, as well as a combination of multiple cloud providers. Nearly three-quarters (72%) of respondents to 451 Research's Voice of the Enterprise: Cloud, Hosting & Managed Services, Workloads & Key Projects 2020 survey stated that they currently have, are implementing or are planning a hybrid IT environment that leverages both on-premises systems and off-premises cloud/hosted resources in an integrated fashion.
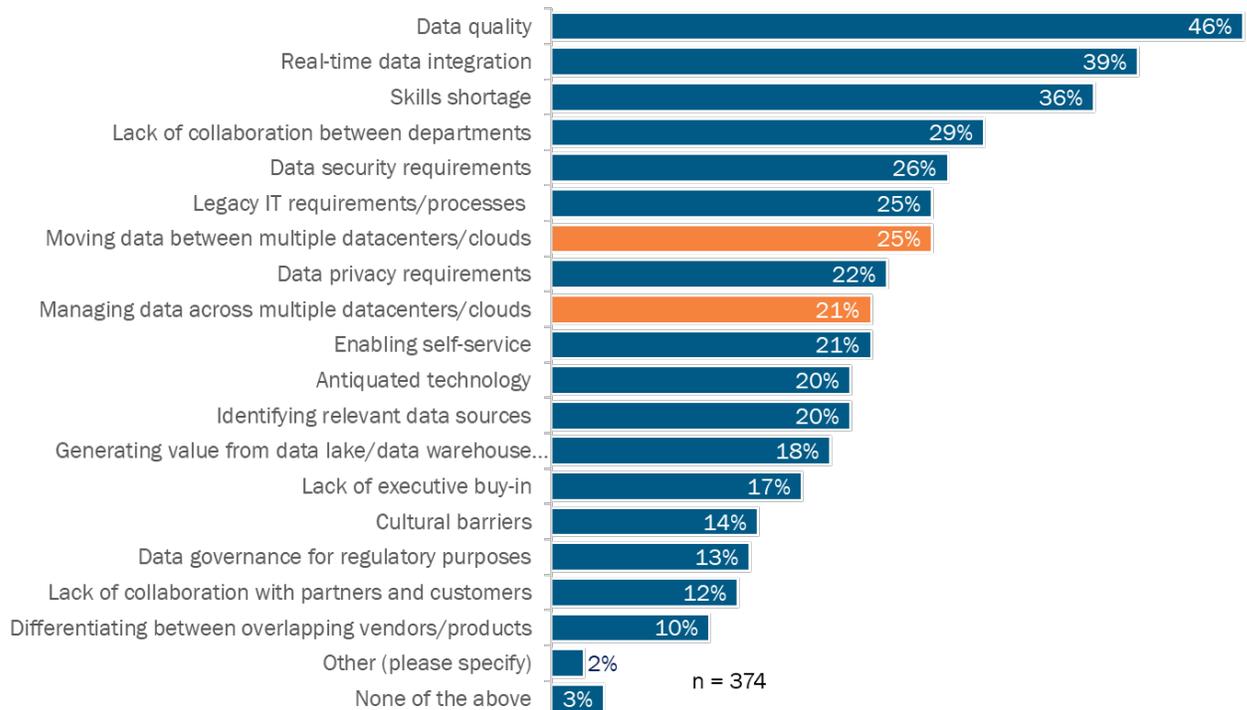
Coincidentally, the same proportion (72%) of companies using the public cloud stated in 451 Research's Voice of the Enterprise: Cloud, Hosting & Managed Services, Workloads & Key Projects 2019 survey that they are using more than one cloud provider. This has significant implications for data processing and analytics. As data is increasingly stored in multiple locations, enterprises need to process and analyze data across those multiple locations. An increase in data processing location choices requires enterprises to make tactical and strategic decisions about when and where to process data, as well as to formulate strategies for data movement or federated data processing and analysis.

451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2020 survey showed that 25% of respondents see moving data between multiple datacenters/clouds as one of the biggest analytics challenges faced by their organization, slightly ahead of managing data across multiple datacenters/clouds (21%).

Biggest Analytics Challenges

Source: 451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2020

Challenge	Percentage
Data quality	46%
Real-time data integration	39%
Skills shortage	36%
Lack of collaboration between departments	29%
Data security requirements	26%
Legacy IT requirements/processes	25%
Moving data between multiple datacenters/clouds	25%
Data privacy requirements	22%
Managing data across multiple datacenters/clouds	21%
Enabling self-service	21%
Antiquated technology	20%
Identifying relevant data sources	20%
Generating value from data lake/data warehouse...	18%
Lack of executive buy-in	17%
Cultural barriers	14%
Data governance for regulatory purposes	13%
Lack of collaboration with partners and customers	12%
Differentiating between overlapping vendors/products	10%
Other (please specify)	2%
None of the above	3%

n = 374

Not surprisingly, 85% of respondents to 451 Research's Voice of the Enterprise: Data & Analytics, Data Platforms 2020 survey agreed that the ability to run the same database on multiple cloud/datacenter environments is an important consideration for their organization when selecting a new data platform.

The ability to use the same database in multiple locations (both on-premises and public cloud) is already addressed by most specialist database providers today, but is relatively nascent in the strategies of the cloud providers themselves. There are examples of on-premises and multicloud offerings from the cloud providers, such as Google BigQuery Omni, Amazon RDS and EMR on AWS Outposts, and Microsoft Azure SQL Managed Instance and Azure Database for PostgreSQL on Azure Arc.

However, for each of the major cloud providers, there are many more examples across their portfolios of data services that are only available on their own cloud platforms. This provides the established data platform vendors a distinct advantage with customers that specifically have hybrid and multicloud support as a strategic requirement.

In many cases, this might be more of a theoretical advantage than a practical one. For many enterprises, the potential freedom to move data between processing locations is more significant than their actual need to do so, and 451 Research's Voice of the Enterprise: Cloud, Hosting & Managed Services, Workloads & Key Projects 2019 survey saw 45% of organizations with more than one cloud provider say that more than 80% of their usage is with their primary vendor.

Nevertheless, there are multiple practical reasons why support for both hybrid IT and multicloud configurations could be considered a genuine advantage. Below are four of the most significant.

On-premises dependency

The results of 451 Research's Voice of the Enterprise: Data & Analytics, Data Platforms 2020 show that 35% of enterprises expect existing on-premises data platform/application stack deployments to remain on-premises, while 33% expect new data platform/application stack deployments to be deployed on-premises.

The primary reasons respondents gave for retaining but modernizing existing data platforms on-premises were data/system security, data locality/sovereignty and leveraging existing IT investments. Application dependency, performance and uptime requirements, staff expertise, and regulatory compliance also have a role to play.

Whatever the reason, a proportion of data workloads will remain on-premises going forward, giving vendors with an on-premises offering an advantage over those that are only available in the cloud. This primarily applies to the workloads in question, which could remain with their existing provider while new workloads are deployed on new databases, either on-premises or in the cloud, but the more mission-critical those existing workloads are, and the longer they can be expected to remain on-premises, the more an existing supplier can leverage them to its strategic advantage.

Geographic requirements

All the major cloud providers can claim to cover most of the globe, but there are some significant gaps, and their regional investments do not completely overlap. For example, AWS and Azure both offer regions in South Africa, but Google Cloud currently does not. In contrast, Google Cloud has a region in Finland, while AWS and Azure do not. Meanwhile, only Azure has announced plans to provide a region in New Zealand, while only AWS currently offers a region in Sweden, although Azure is coming soon.

These geographic 'dead spots' are not going to impact every company, but for multinational companies, the fact that there are some counties for which a choice of cloud providers is not available could be problematic. This issue could require them to maintain on-premises resources or take advantage of multiple cloud providers, in order to deliver the required performance or support data sovereignty requirements in individual regions. Data locality/sovereignty is the second-most-popular reason for modernizing data platform workloads on-premises (rather than moving them to the cloud) according to 451 Research's Voice of the Enterprise: Data & Analytics, Data Platforms 2020.

Migration to or across clouds

Even if a company is planning to move all or most of its existing on-premises database workloads to the cloud, hybrid IT support is arguably an advantage in facilitating modernization and migration to the cloud. For example, Cloudera can cite the delivery of CDP Private Cloud as an important offering in terms of enabling existing customers to modernize their existing on-premises Cloudera Enterprise Data Hub or Hortonworks Data Platform deployments. While some will remain on-premises, CDP Private Cloud can also be used as a stepping stone to facilitate migration to CDP Private Cloud.

Support for multiple clouds doesn't necessarily mean support for the ability to easily replicate or move data between clouds. Take, for example, MongoDB Atlas. Although it provides users with a choice of AWS, Google Cloud or Microsoft Azure, once that choice is made, it cannot be easily changed. Moving Atlas workloads between clouds involves manually complex data extraction and migration. At least it did until the delivery of MongoDB Atlas multicloud clusters, which enable data to be distributed across AWS, Google Cloud and Azure simultaneously, with the ability to add, remove or migrate across clouds as a configuration option. This functionality could potentially be used to facilitate migration between cloud providers.

Replication across clouds

Replicating data between clouds also enables users to potentially take advantage of different cloud-specific services for different purposes. For example, the operational application and its associated database may be running on one cloud provider, but the data could be automatically replicated to the same database running on a different cloud to take advantage of analytics capabilities that are only available from the second cloud provider.

This is a simplified theoretical example, but our surveys tell us it is playing out in practice. Access to vendor-specific platform capabilities is the number one reason for using multiple IaaS/public cloud vendors, according to 451 Research's Voice of the Enterprise: Cloud, Hosting & Managed Services, Workloads & Key Projects 2019.

Another potential reason for replicating data across cloud providers is high availability – providing insurance against cloud service disruption. While each cloud provider offers multiple regions and availability zones to support resiliency, some data cannot be moved between cloud regions for regulatory reasons. Replicating data between multiple cloud providers, or even between on-premises and cloud, within a region is a potential solution to mitigate risk while complying with data residency requirements.