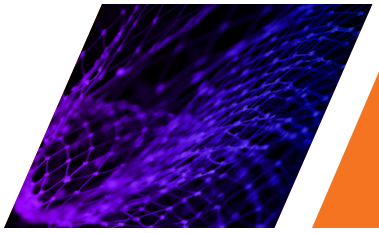
June 2020

How Data Mobility Unlocks Public Cloud Cost Savings

In this issue



How Data Mobility Unlocks Public Cloud Cost Savings

According to Gartner, more than 80% of all organizations are using more than one public cloud and are building applications and products that require intercloud data mobility. Regardless of where the data sits, it is important that the underlying cloud resources deliver the needed performance and capacity, with the ability to scale out in a cost-efficient manner. In addition, companies need specific tools to orchestrate and monitor intercloud data mobility in order to support fast-changing business circumstances.

Silk

Silk's Cloud Data Platform was designed and built for the public cloud. Our software-only architecture provides a single, virtualized, multi-cloud data layer which delivers industry-leading consistent performance for business-critical applications. In addition, Silk's enterprise-class data services combined with machine learning-based analytics and autonomous orchestration provides a cost-efficient approach to support intercloud data mobility. Silk can help your organization run business-critical applications in cloud-agnostic environments, cutting your cloud spend by 30%.

Silk + TPG Software

TPG Software, a provider of investment accounting software solutions, was switching to a SaaS delivery model in order to stay competitive and provide a consistently good customer experience. Yet this move presented a few challenges. TPG customers generate hundreds of reports with peak activity taking place during the middle and end of each month. To accommodate these peak workloads, TPG's infrastructure would need to have the flexibility to add more resources when needed and then turn these resources off during slower times.

With the Silk Cloud Data Platform, TPG is able to scale their resources in the public cloud. This agility means TPG can run at full efficiency and minimize the costs of managing its infrastructure by 30%. Minimizing costs even further, Kaminario enabled TPG to securely consolidate several application servers to fewer SQL servers. This resulted in over 10x savings on database license costs versus what the company would have spent if it had self-deployed the application servers in its own datacenter. With the ability to scale resources up and down to meet customer demands, Silk gave TPG extreme flexibility at the most cost-efficient price point.

Source: Silk

Research from Gartner:

Are You Ready for Multicloud and Intercloud Data Management?

The increased adoption of multicloud and intercloud deployments in support of data management solutions has important implications for data and analytics strategies. Data and analytics leaders must prepare for these impacts now, to ensure optimal use of cloud resources.

Impacts

- A recent Gartner survey on cloud adoption revealed that 80% of respondents using public cloud are using more than one cloud service provider (CSP). These multicloud architectures often arise organically, through the consumption of specific services or SaaS offerings that may not align with the primary cloud strategy, and therefore add complexity to cloud operations.

- CSP offerings are becoming the new platform, and have all of the integration challenges that on-premises platforms had before them. When used in multicloud and intercloud scenarios, there is the added complexity of network latency, as well as data transfer concerns.

- Independent software vendors (ISVs) and containers provide cloud-agnostic solutions, which challenge some closed CSP services.

Recommendations

For data and analytics leaders planning to incorporate multiple CSPs into their data management solutions:

- Do not try to restrict the use of multiple clouds; it will be a losing battle. Rather, create a policy around cloud use and adoption — for tracking cloud spend and usage across multiple clouds — to get an understanding of how cloud resources are being used and to prevent uncontrolled spending and the use of nonstrategic cloud offerings.
- Implement tooling to track usage across clouds — for budgeting and resource allocations and latency — to identify architectural pain points that may need remediation.
- Select a solution that extends the cloud object store to multiple clouds, as a default multicloud deployment tier, because it will allow for the greatest degree of flexibility. Using the application or DBMS tier may be better-suited to addressing specific needs, but will be less flexible.
- Evaluate ISV offerings that extend the breadth and scope of what is available in a native CSP offering, if you are considering active integration of data across multiple clouds.

Strategic Planning Assumptions

By 2023, 75% of all databases will be on a cloud platform, reducing the DBMS vendor landscape and increasing complexity for data governance and integration.

By 2022, 50% of cloud buying decisions will be based on data assets provided by the cloud service provider rather than on product capabilities.

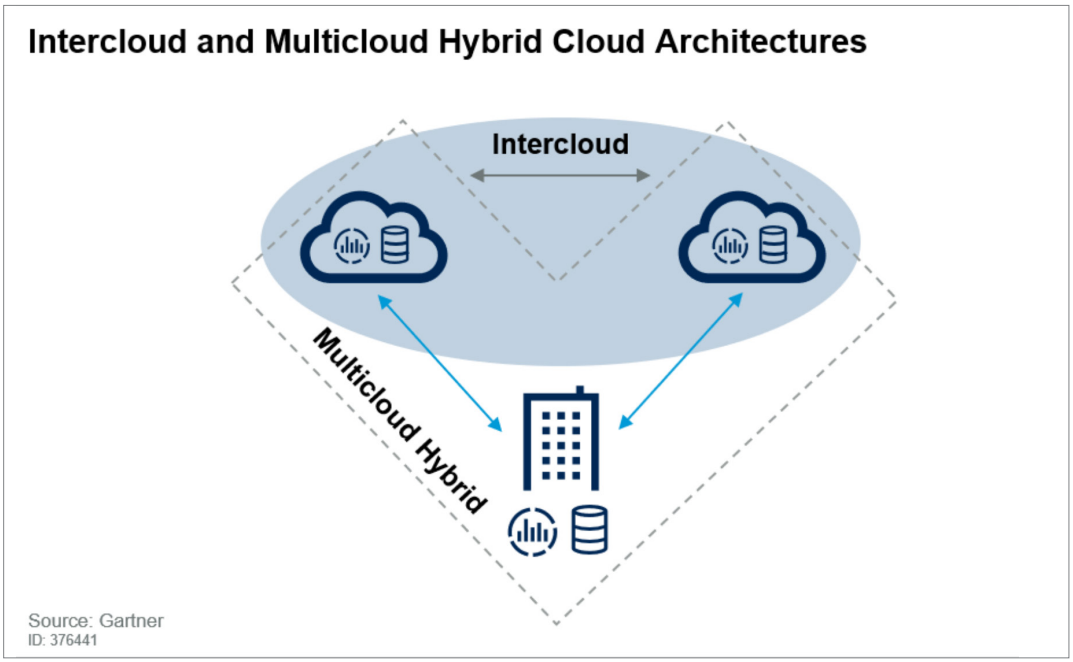
Analysis

Multicloud and intercloud data management deployments are here (see Note 1). Gartner inquiry data shows a nearly 300% increase in inquiries mentioning "multicloud" in 2018, compared with 2017.[1] Further, a recent Gartner cloud adoption survey, conducted in October and November 2018, found that 80% of those using public cloud were using more than one cloud service provider (CSP).[2]

All of the complexities of architecture spanning hybrid cloud and use-case-specific hybrid cloud apply to multicloud and intercloud deployments, with some additional complexities.

It is important to make a distinction between "multicloud," which simply refers to similar applications being deployed on multiple cloud environments, and "intercloud," which involves data transfer between clouds (see Figure 1).

Figure 1. Intercloud and Multicloud Hybrid Cloud Architectures



Data and analytics leaders must prepare for the complexities of multicloud and intercloud deployments, to avoid potential performance issues associated with data latency, unplanned cost overruns and the ongoing difficulties associated with integration efforts, data transfer fees and increased complexity. Figure 2 summarizes the impacts and our top recommendations for data and analytics leaders involved in preparing for this change.

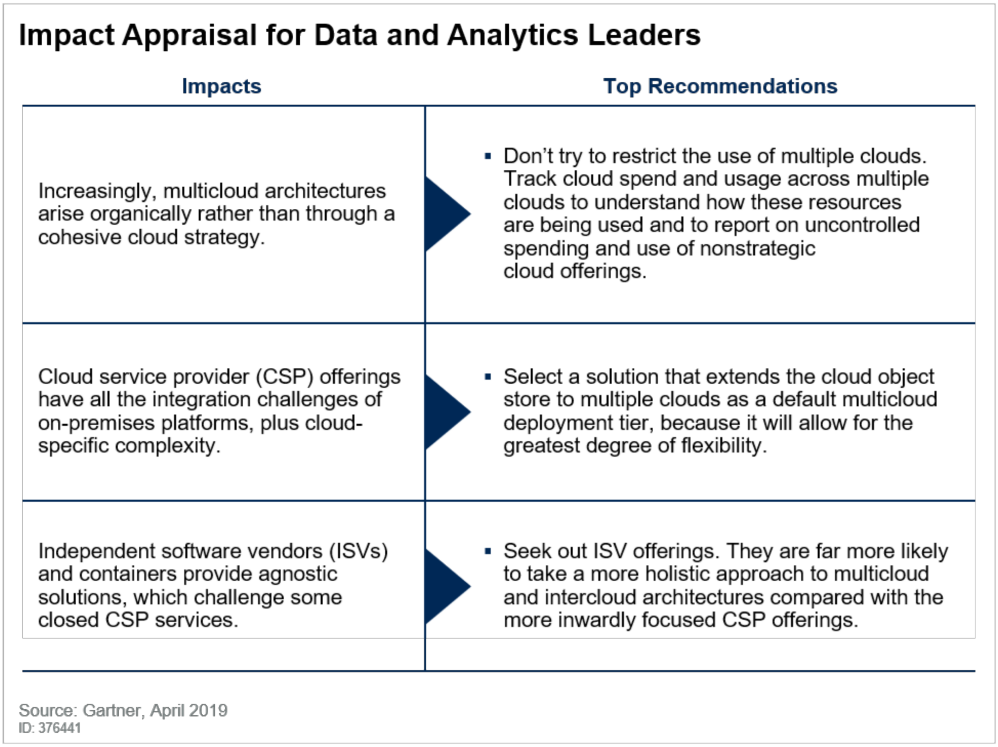
Impacts and Recommendations

Increasingly, Multicloud Architectures Arise Organically Rather Than Through a Cohesive Cloud Strategy

Data and analytics leaders have entered a new era of complexity as multicloud and intercloud architectures are becoming common (see Note 1). While many organizations strategically align with a single cloud provider, it is nearly impossible to enforce strict adherence to a single provider due to the following factors:

- Cloud services are easy to deploy and often disregard traditional IT budgeting, being done most often via discretionary operational budgets owned by lines of business, rather than centralized IT.
- Specific advantages or capabilities from point solutions on additional, nonprimary CSPs are commonly available.
- Many organizations deploy software as a service (SaaS) offerings that represent yet another potentially closed cloud environment, with separate cloud infrastructure and data persistence tiers.
- Organizations may decide to use development or visualization tools on one CSP, with data originating or residing on another.

Figure 2. Impacts and Top Recommendations for Data and Analytics Leaders



As soon as any part of the organization begins to use cloud services in a second CSP, and to store data there, that is a multicloud environment and the organization will, potentially, have to deal with intercloud data management. There are specific concerns around intercloud data integration needs. These range from latency through governance concerns (including security, encryption and data sovereignty), and even mundane concepts such as the likelihood of potential data transfer fees that are inherent to intercloud distributed data management.

Data and analytics leaders must accept this new reality, but can also create policies that mitigate risk and complexity. In other words, they should embrace multicloud deployments when the benefits outweigh the added complexity or compromises required, and follow a "priority primary cloud" strategy to make sure practitioners give adequate consideration to offerings in their primary, strategic cloud. The use of a Tier 1/ Tier 2 nomenclature — similar to that commonly used to set policy around DBMSs within an organization — can provide common guidance. By default, any new initiative should leverage Tier 1 categorized cloud assets. With Tier 2 services, any initiatives are undertaken at your own risk, and with your own budget.

Recommendations for data and analytics leaders:

- Do not try to restrict the use of multiple clouds; it will be a losing battle. Rather, create a policy around cloud use and adoption to track cloud spend and usage across multiple clouds. This will lead to an understanding of how cloud resources are being used, and to reporting on uncontrolled spending and use of nonstrategic cloud offerings.
- Follow a preferred primary cloud provider strategy and require justification for use of nonstrategic cloud assets, categorizing them with Tier 1 and

Tier 2 offerings and setting default standards for new projects.

- Put in place robust monitoring tools that will manage resource consumption across multiple clouds and reduce cost overruns. Single-pane-of-glass monitoring (presenting data from multiple sources in a unified display) may also be used to ensure data governance and security policies are being adhered to.

CSP Offerings Have All the Integration Challenges of On-Premises Platforms, Plus Cloud-Specific Complexity

The CSP infrastructures and the services that run on them are becoming the new data management platform. These platforms consist of:

- General infrastructure that supports infrastructure as a service (IaaS) initiatives
- Common service layers, such as a cloud object store that can be used as a data management fabric spanning cloud services
- Native CSP point solutions designed to support specific needs through a best-fit engineering approach (see Note 2)
- Third-party independent software vendor (ISV) point solutions that run on cloud infrastructure

A sample listing of some CSP-native and some ISV cloud offerings for data management is shown in Figure 3 and Figure 4.

Figure 3. Comparison of Sample CSP and ISV Cloud Offerings

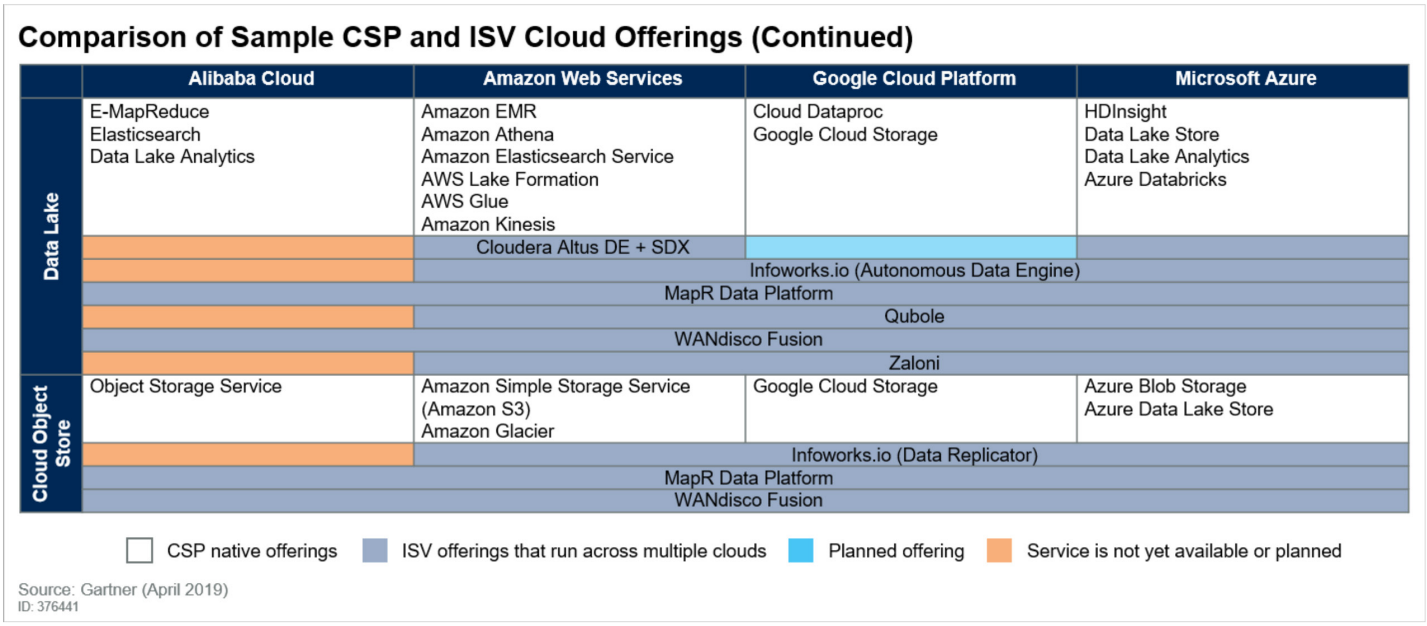
Comparison of Sample CSP and ISV Cloud Offerings

	Alibaba Cloud	Amazon Web Services	Google Cloud Platform	Microsoft Azure
OPDBMS	ApsaraDB for MariaDB TX ApsaraDB RDS (MySQL, SQL Server, PostgreSQL, PPAS) POLARDB Distribute Relational Database Service	Amazon Aurora (MySQL & PostgreSQL compatible) Amazon Relational Database Service (Amazon RDS) (SQL Server, Oracle, MySQL, Postgres, MariaDB) Amazon RDS on Vmware	CloudSQL (MySQL, Postgres, SQL Server) Cloud Spanner	Azure SQL Database Azure Database for MySQL Azure Database for PostgreSQL Azure Database for MariaDB SQL Server on Virtual Machines
		Datastax Constellation/Datastax Managed Services		
		MapR Data Platform		
		MarkLogic Data Hub Service		
Data Warehouse	HybridDB for PostgreSQL MaxCompute DataWorks	Amazon Redshift	BigQuery	Azure SQL Data Warehouse
		Cloudera Altus DW		
		MapR Data Platform		
		MarkLogic Data Hub Service		
		Snowflake Computing		
		Teradata Vantage		
Nonrelational DBMS	ApsaraDB for MongoDB ApsaraDB for Redis ApsaraDB for Memcache Time Series Database Table Store	Amazon DynamoDB Amazon ElastiCache Amazon Neptune Amazon Timestream Amazon DocumentDB (with MongoDB compatibility) Amazon Quantum Ledger Database (QLDB)	Cloud Bigtable Cloud Firestore	CosmosDB Table Storage Redis Cache HBase in HDInsight
		Datastax Constellation/Datastax Managed Services		
		MapR Data Platform		
		MarkLogic Data Hub Service		

CSP native offerings | ISV offerings that run across multiple clouds | Planned offering | Service is not yet available or planned

Source: Gartner (April 2019)
ID: 376441

Figure 4. Comparison of Sample CSP and ISV Cloud Offerings (Continued)



Just as there were integration challenges with disparate on-premises platforms, so there are similar integration challenges in the cloud; but the cloud integration challenges are multidimensional, consisting of:

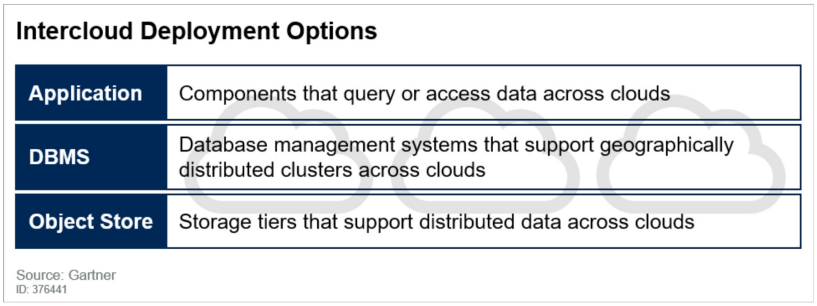
- **Single cloud service integration** — This involves integrating applications and data management among various services within a single CSP. The cloud object store often serves as a common data fabric, but different CSPs will provide different levels of cohesiveness between their various offerings.

- **Intercloud integration (native CSP offerings)** — This has all the challenges that on-premises integration has, with the added complexity of data latency, data transfer (and associated fees), different cloud provisioning tools and capabilities, different financial governance controls, and potential regulatory concerns associated with data movement between clouds. These offerings are generally concerned with moving data into their specific CSP infrastructure, as opposed to between multiple CSPs.

- **Intercloud integration (ISV offerings)** — Most of these tools have significantly mature offerings regarding metadata, data quality and design assistance. The current challenge is in adding administrative and management capabilities, that pass operational and runtime parameters, to CSP environments to allocate infrastructure resources on the cloud provider — which differ on the various CSPs. ISV data integration providers have long experience in adapting to new platforms, operating systems and data management platforms.

Further, in an intercloud world, the client endpoints of integration can fall within three components in the data management stack (see Figure 5).

Figure 5. Intercloud Deployment Options



At the application layer, we include components that query data or access data across clouds. Some examples of this include data virtualization tools such as Denodo, or BI query engines such as Microsoft PowerBI. The biggest concern with intercloud data management at the application layer is likely to be data latency and performance.

At the DBMS layer, we include those data management platforms that can implement geodistributed clusters spanning more than one cloud. Examples include Apache Cassandra, CockroachDB and NuoDB. The primary concern with intercloud data management at the DBMS layer is likely to be latency — for data that spans multiple clouds — because of the distances between physical data centers. As always, there is a concern with data consistency when using intercloud nonrelational database services that do not support strong consistency.

At the object store layer, we include those vendors that are working on managing data distribution and replication across clouds. Many of the legacy Hadoop vendors fall into this category, as well as new vendors such as WANdisco. Ideally, these solutions should provide the ability to govern which data gets replicated and to where, and optimize by only moving data that has changed — to avoid unnecessary egress costs.

Considering that the object store within a single cloud deployment often becomes the data clearing house that ties disparate services together, extending this concept across clouds can be a powerful approach to a unified multicloud and intercloud data management environment.

Recommendations for data and analytics leaders:

- Select a solution that extends the cloud object store to multiple clouds as a default multicloud deployment tier, because it will allow for the greatest degree of flexibility. Using the application or DBMS tier may be better-suited to address specific needs, but will be less flexible.
- Do not underestimate the added complexity of running multiple cloud environments; a single set of policies may need to be implemented multiple times with differing capabilities and tools.

ISVs and Containers Provide Cloud-Agnostic Solutions and Challenge Some Closed CSP Services

Multicloud ISV offerings running on cloud infrastructure promise cloud agnosticism and a means of countering any fears of cloud vendor lock-in by offering a choice of CSP platform. Migrating from one

DBMS to another is time-consuming and resource-intensive. If we can eliminate the effort to rewrite an application, and only focus on extracting data from one cloud to another, the task is radically simplified.

Containers provide a layer of abstraction from underlying cloud or on-premises infrastructure, allowing applications to move as a single unit between container management services. Given that containers are stateless, and data management is inherently stateful, a common data persistence tier needs to be provisioned for data-centric applications. Any container service should be able to work with both on-premises and cloud-native data persistence tiers (adding an additional level of flexibility).

Both of these approaches face long-term challenges from native CSP offerings. Both will need to highlight their existing differentiation, relative to native CSP offerings, in order to secure their existing competitive advantages. This differentiation will focus on two primary dimensions:

- The multicloud promise of portability (and the possibility of running an architecture spanning an intercloud environment).
- Differentiated core product offerings that are sufficiently different from native CSP offerings.

The promises of cloud portability are, in most cases, likely to be oversold. Moving from one CSP to another has far greater implications than the single component that a cloud-agnostic vendor is likely to be able to address. If a multicloud portability strategy provides significant business value over and above what the native CSP offering provides, it may make sense. This scenario is likely to be a minority occurrence, however, and largely focused on those parts of the stack that are already concerned with platform flexibility and agnosticism; for example, the data integration vendors.

Further, as native CSP offerings become more mature, with greater capabilities, the long-term prospects of maintaining the second of these two points of differentiation will be increasingly difficult. Especially as CSPs take advantage of tight integration with an underlying infrastructure that may provide operational efficiencies and may not be available to ISV offerings.

Recommendations for data and analytics leaders:

- Evaluate ISV offerings that extend the breadth and scope of what is available in a native CSP offering if you are considering active integration of data across multiple clouds.
- Expect ISV offerings to offer significant differences in capabilities in those areas, such as data integration, that are specifically focused on distributed data management.

Acronym Key and Glossary Terms

CSP	cloud service provider
ISV	independent software vendor

Evidence

[1] Gartner Inquiry Service — We examined more than 700 inquiries taken across Gartner's inquiry service for mentions of "multicloud" or "multi-cloud."

[2] Gartner's Cloud Study 2018 (P-18029 Cloud Adoption). The survey was conducted online by an external partner, between October and November 2018. The full study surveyed 1,200 individuals, of which 628 respondents reported that their organization was using the public cloud, and 507 reported using more than one public cloud provider. Results of this study do not represent "global"

findings, or the market as a whole, but are a simple average of results for the targeted countries covered in this survey.

Note 1

Hybrid Cloud, Multicloud and Intercloud Taxonomy

There is a lack of clarity in the market around the definition and meaning of these terms. Vendors will often refer to their product as "multicloud" — by which they mean it runs on more than one cloud. Gartner accepts this definition, but we find it incomplete. Below, we articulate a full taxonomy for hybrid cloud, multicloud and intercloud architectures.

- **Hybrid Cloud** — This refers to implementations that span on-premises and cloud deployments. In "3 Ways That Hybrid Cloud for DBMS Will Drive Your Data Management Strategy," we further define the subcases of:
  - **Architecture Spanning Hybrid Cloud** — Where components of a single logical deployment span on-premises and cloud.
  - **Use-Case-Specific Hybrid Cloud** — Where different components are segmented by their development life cycle function (for example, development, test, production).
- **Multicloud** — A service or product runs on more than one cloud service provider infrastructure, and may also run on-premises.
- **Intercloud** — Where data is integrated or exchanged between cloud service providers as part of a logical application deployment.

These definitions can also be combined; for example, "architecture spanning intercloud" would refer to a logical deployment of a single application — where components were deployed on more than one cloud and regularly exchanged data between the clouds.

Similarly, "use-case-specific multicloud" would mean that a development environment was deployed in one cloud and a production environment in another.

You could not have "use-case-specific intercloud," because data is not exchanged actively between the environments in a use-case-specific scenario.

Note 2

What Is Best-Fit Engineering?

Best-fit engineering refers to the deployment of specific solutions to address targeted needs within the data management environment. It is common in the cloud, where a CSP portfolio consists of a broad range of services that perform a specific task and make up a larger ecosystem. While the cloud is characterized by best-fit point solutions, the cloud ecosystem and infrastructure can serve as the basis for a platform play. This applies to native CSP product offerings as well as to point solutions from ISVs.

Source: Gartner Research, G00 376441, Adam Ronthal, Rick Greenwald, Donald Feinberg, 24 May 2019

Silk

Contact us

For more information contact us at: