Death Of The Cloud Storage Gateway
Death Of The Cloud Gateway

Hybrid cloud storage has become one of the latest hot storage market trends. The concept is to blend on-prem storage with off-site public cloud object storage or on-site private cloud object storage. Hybrid cloud storage conceptually integrates two distinctly separate and frequently geographically dispersed storage systems making them appear as one. The motivation behind hybrid cloud storage is to seamlessly provide the performance, security, and control of local on-prem primary storage with the unlimited scalability, flexibility, durability, elasticity, and pay-as-you consume cloud economics of cloud object storage. Hybrid cloud storage can have tremendous appeal depending on how it is architected for the storage problems it solves. There are four common hybrid cloud storage implementations or “models” in use today. Each one can claim to be hybrid cloud storage even though they are functionally quite different.

Four Most Common Hybrid Cloud Storage Models In Use Today

The hybrid cloud storage models are:

1. Cloud integrated storage (CIS) a.k.a. cloud storage gateway or appliance
2. Storage tiering to the public cloud
3. Storage system to virtual storage appliance (VSA) running in the cloud
4. Private on-prem object storage to public cloud object storage

Cloud Integrated Storage (CIS) a.k.a. Cloud Storage Gateway Or Appliance

Cloud Gateways represent the oldest hybrid cloud storage model. To date, there have been no successful exits for any standalone Cloud Gateway vendor. They have all either collapsed—such as Cirtas with its much hyped “BlueJet” gateway—or sold as a “technology tuck-in,” such as TwinStrata did when it sold its “CloudArray” to EMC. The standalone cloud gateway providers that remain—Nasuni, Panzura, Ctera, among others—are basically middle men, intermediaries in a world that is quickly absorbing their purported value prop into other more strategic assets in IT environments. Established vendors with cloud gateways, such as NetApp with AltaVault, aren’t in a much better position as they charge $1M for a typical enterprise configuration—very pricey for an intermediary device that adds more cost and complexity.

The functionality of a cloud gateway can be summarized as cache locally on prem, store in the public cloud. The CIS is integrated with the public storage cloud via the RESTful API generally the S3 standard. Data is written to the CIS, deduplicated and compressed, so as to consume as little public cloud storage as possible and save cloud storage fees. The data sent to the public cloud storage leaves a stub in the CIS. All data sent to the CIS is presented as being local in the CIS. In reality, the vast majority of that data is a stub that when accessed recalls the actual data from the off-site public cloud storage. The CIS stays populated with only the most frequently accessed data as a read cache to reduce the lengthy latencies that occur from accessing data offsite across the Internet. Some CIS implementations utilize first-in-first-out (FIFO) read caching where older data is automatically expelled, as newer data is input to cache.

There are several stark problems with this model. The biggest is that most CIS systems are simply not designed to be primary storage. They typically lack primary storage performance and/or storage services such as snapshots, writable clones, replication, on-prem scalability, block and file protocol unified...
support, replication, and much more. Many are just a single controller, in other words, not HA (highly available) and not ALUA (asymmetrical logical unit access). This makes the CIS a secondary storage system. It sits as an intermediary between primary storage and cloud storage. Data on the primary storage has to be moved or migrated in some fashion to the CIS to utilize cloud storage. That migration requires a data mover.

The data mover is often pull-based such as homegrown scripting, open source, (both of which require a lot of manual effort), or third party archiving or backup software. The software identifies data that needs to be stored in the cloud based on user policies and migrates it to the CIS. Note that most of the third party software can also move the data directly to cloud storage, bypassing the need for a CIS altogether. The cost of data mover software, plus the cost of the servers that software runs in, plus the cost of CIS can be considered a cloud storage access tax making the cloud storage prohibitively expensive.

The less prevalent CIS option is for data to be pushed from the primary storage system to the CIS. Few primary storage systems can do this because doing so requires integration with the CIS. The primary storage is essentially replicating to the CIS. More often than not it requires both the primary storage and CIS to come from the same vendor. Once again, the cost of the primary storage replication software license plus the CIS adds an extortionately expensive cloud storage access tax.

There’s a visibility and troubleshooting problem with CIS hybrid cloud storage. The CIS is a black box. Visibility stops at its edge. Primary storage cannot see the cloud storage behind it. There is no end-to-end visibility or analytics. Troubleshooting is complicated at best or impossible at worst.

There is one other significant issue with CIS hybrid cloud storage. Data moved through the CIS to cloud storage can only be read back through the CIS. And since most CIS are single controllers, it is highly recommended that there be at least two CIS systems on-prem and one more somewhere else for business continuity and disaster recovery. That adds even more cost or cloud storage access taxes to this hybrid storage cloud. Data that passes through a Cloud Gateway is held hostage by that gateway and can’t be read unless it’s retrieved by the gateway from the cloud and then read on the gateway. All of this is now integrated into other solutions so customers don’t need to practice their systems integrator skills.

**Storage Tiering To The Public Cloud**

A more recent hybrid cloud storage model is the ability of some primary storage systems to treat public cloud storage as a storage tier. It is a modern implementation of hierarchical storage management (HSM) with cloud storage being the target storage tier. It eliminates an intermediary to cloud storage. The primary storage system migrates data to the public cloud based on user-defined policies based on data age, value, time since last access, and more. After moving the data, it creates a stub that makes it appear that the data is still where it originally landed in the storage system. Reading the data takes a bit longer as the data is recalled from the cloud storage and rewritten back to its original storage tier. If the primary storage deduplicates and compresses the data, it won’t rehydrate the data before it’s moved. This hybrid cloud storage model too has significant issues.

Similar to the CIS hybrid cloud storage, it prevents the primary storage system from utilizing the cloud storage as a directly accessible storage pool. It can only utilize it as an HSM target. This is primarily because the cloud connector is an internal software cloud storage gateway. It too lacks end-to-end...
visibility and analytics. Most of the primary storage vendors with this capability charge a licensing fee per TB to move data to cloud storage thereby imposing an exorbitant cloud storage access tax.

Storage System To Virtual Storage Appliance (VSA) Running In The Cloud

This hybrid cloud storage model is a variation of storage-to-storage replication. The target storage is a virtual storage appliance (VSA) variation of the primary storage. The VSA runs as a VM instance in the public cloud compute. That instance is directly connected to the cloud storage of that public cloud provider. Data is replicated from the primary storage to the target VSA where it is then directly stored in the public cloud storage. The real value here is business continuity and disaster recovery. It reduces their cost through utilization of public clouds as the business continuity and DR data center. Snapshots and incremental snapshots are replicated from the primary storage to the VSA creating a duplicate set of data at both locations while also protecting the data from hardware and site failures, human error, malware, and maliciousness. It also delivers value for data sharing between applications running on-prem and applications running in the public cloud. But primary storage system to VSA has other substantial issues. Data moved between the primary storage and cloud-based VSA is duplicated. That eliminates it as a good or quality archival solution. It consumes too much cloud storage, swelling the cloud storage fees. Once again there is an intermediary between the primary storage and the target cloud storage making end-to-end visibility, analytics, and troubleshooting much more complicated. And the cloud storage access taxes are evident again with the replication license fees as well as the VSA license fees (it is highly recommended to deploy at least two VSAs for high availability) plus the cloud compute instance fees for the VSAs. It makes the cloud storage costs unreasonable.

Private On Prem Object Storage To Public Cloud Object Storage

The object storage to object storage hybrid cloud has multiple variations. The first is when the on-prem object storage is the same as that in the public cloud. In this variation the public cloud storage is an extension or remote nodes of the on-prem object storage. It is one object storage system. It functions as a geographically distributed scale-out object storage system. All nodes both on prem and in the cloud have access to all other nodes. This is a very simple hybrid cloud storage that makes the cloud storage indistinguishable from the on-prem storage. Data access policies are essential in making sure data is placed in such a way that Internet latencies do not radically slow down applications.

In the second variation, the on-prem object storage is different from the object storage in the cloud. The two object storage systems are unique and distinct from one another. They are two separate systems. Data must be copied from the on-prem object storage to the public cloud storage typically on the S3 RESTful API. It is storage-to-storage data transfers with many of the same issues as with other storage-to-storage data transfers such as duplication of consumed capacity, higher overall costs, etc.
Both of these object-to-object storage hybrid cloud storage variations are for all intents and purposes secondary storage. Object storage by its nature has high latencies before the media (HDD or SSD) are even considered. This makes it mostly unsuitable as primary storage, but great as a target for archiving, backup, and stale data long-term storage. It has a relatively low cost, but it's not what IT organizations mean when they are talking about hybrid cloud storage. They mean primary storage on-prem to public cloud object storage. For the vast majority of organizational storage use cases, this is a no-op.

In contrast, a hybrid cloud storage implementation ideally behaves as:

- A contiguous homogeneous storage
- Where the on-prem primary storage transparently sees, manages, and uses the public cloud storage as an addressable storage tier

And this is where Oracle’s Cloud Converged Storage is completely changing the game and making the standalone cloud gateway vendors—and large storage vendors peddling their own in-house gateways—search for new business models in the hopes of maintaining relevancy in an increasingly changing IT landscape.

**Oracle Cloud Converged Storage**

Oracle Cloud Converged Storage deeply integrates Oracle Storage Cloud Services with high-performance Oracle ZFS Storage Appliances. It integrates at the OS layer using Oracle ZFS Cloud software and makes the Oracle Storage Cloud a transparent storage pool to the Oracle ZFS Storage Appliances. Oracle Cloud Converged Storage makes Oracle Cloud Storage Services look, feel, and act like any other ZFS Storage Appliance native HDD or flash SSD storage pool. It works with both Oracle Storage Cloud Services – Object Storage and Archive Storage. Oracle is first to the market with Cloud Converged Storage. There are no intermediary systems, gateways, third-party data mover software, and most importantly no cloud storage access fees or taxes. In fact, this capability is included in the latest Oracle ZFS OS upgrade, as a no-cost option. There has never been anything like it until now.

What Cloud Converged Storage does is change how cloud storage is efficiently and cost effectively utilized. For example:

- Oracle Storage Cloud can be directly assigned to specific application workloads as storage pools within the ZFS Storage Appliance. Application workloads that do not need high performance and can work with higher degrees of latency can directly read and write through the ZFS Storage Appliance to the Oracle Storage Cloud, reducing on-prem capacity requirements.

- Passive or stale data on the ZFS Storage Appliance can be archived and migrated to the Oracle Storage Cloud storage pool automatically while leaving a stub in its original location. Since most (85 to 90%) data center data is stale, rapidly becoming stale ninety days after creation, this capability measurably reduces on-prem capacity requirements.

- Snapshots by the ZFS Storage Appliance can be auto-replicated to the Oracle Storage Cloud simplifying disaster recovery (DR) and business continuity (BC). Any of the customer’s ZFS Storage Appliances with the proper credentials can mount that data in the Oracle Storage Cloud immediately. This also simplifies data sharing and workflow collaboration.

- All data on the ZFS Storage Appliance is inline deduplicated and compressed, and the data moved to the Oracle Storage Cloud does not need to be rehydrated before it is moved.

- The ZFS Storage Appliance has the same RESTful object API as Oracle Storage Cloud services in addition to NFS, SMB file interfaces as well as Fibre Channel and iSCSI. This common object API empowers application development to be written once for on-prem and still be capable of running in the Oracle Cloud without change. This capability saves time when moving to the cloud. More importantly it
permits DR directly in the Oracle Cloud and dev ops in the Oracle Cloud when the application will be deployed on prem utilizing the ZFS Storage Appliance with no changes, tweaks or rewrites of the application.

- Security is managed on-prem with the ZFS Storage Appliances and automatically extends to the Oracle Storage Cloud. Simplifying cloud storage security and maintaining on-prem control.

- There are no scalability limits to how much ZFS Storage Appliance data can be put in the Oracle Storage Cloud. This entirely changes storage economics. Much reduced on-prem capacity equals much lower capital expenditures (CapEx). Reduced on prem capacity also reduces on-prem operating expenditures (OpEx). Oracle Storage Cloud capacities are reduced since data is deduplicated and compressed. The Oracle Storage Cloud fees for the actual cloud object storage consumed are on a true pay-as-you go subscription basis paid in arrear. Data storage costs will shift to the Oracle Storage Cloud as the percentage of consumed storage grows in the cloud over time. The remaining data on-prem in the ZFS Storage Appliances should fairly consistently and stable. Reduced on-prem storage capacity leads to lower ongoing tech refresh costs. Lower CapEx and lower OpEx equals a much lower total cost of ownership (TCO) compared to other vendors’ offerings which either require the purchase of hardware (gateways or virtual appliances) or software (data mover or backup software) intermediaries or charge a per terabyte licensing fee—AKA “tax”—to enter the cloud. In contrast, with Oracle ZFS Storage Appliance, the Oracle ZFS Cloud software is included with the OS upgrade, there is no need for intermediaries, and there are no cloud entrance licensing fees.

**Summary and Conclusion**

Hybrid cloud storage has generally held the promise of providing the best of on-premise storage with the best of cloud storage with more control and a lower TCO. Until now, it has fallen far short of its promise. Oracle Cloud Converged Storage is the first to change this game and finally fulfill the promise of hybrid cloud storage. By doing so, Oracle eliminates intermediaries such as the standalone cloud gateway model, and the entire concept of storage access taxes that can more than double the cost of cloud storage. IT organizations are smart and getting smarter over outrageous fees that deliver no value.

**For more information**

Web: [Oracle ZFS Storage Appliance](#)  
Solution Brief sponsored by Oracle. About the author: Marc Staimer, as President of the 19-year-old Dragon Slayer Consulting in Beaverton, OR, is well known for his in-depth and keen understanding of user problems, especially with storage, networking, applications, and virtualization. Marc has published thousands of technology articles and tips from the user perspective for internationally renowned online trades including many of TechTarget’s Searchxxx.com websites as well as Network Computing. Marc has additionally delivered hundreds of white papers, webinars, and seminars to many well-known industry giants such as: Brocade, Cisco, DELL, EMC, Emulex (Avago), HDS, HP, LSI (Avago), Mellanox, NEC, NetApp, Oracle, QLogic, SanDisk; as well as smaller, less well-known vendors/startups including: Asigra, Clustrix, Condusiv, DH2i, Diablo, FalconStor, Gridstore, Nexenta, Neupert, NetEx, NoviFlow, Permabit, Qumulo, StorONE, Tegile, and many more. His speaking engagements are always well attended, often standing room only, because of the pragmatic, immediately useful information provided. Marc can be reached at marcstaimer@me.com, (503)-579-3763, in Beaverton OR, 97007.