Considering On-Premises and Cloud Storage

Sometimes overlooked when planning for whether to move or deploy an application on a public cloud or on-premises are the considerations for storing and managing information. These can be critical to success and mitigation of risks for an organization. While the importance of different factors to consider may vary depending on different organizations/companies, there are a number of important ones that need to be evaluated to make an informed and defensible decision.

A major failing is to not look at storing and managing information on-premises or in a public cloud from a TCO perspective over a set period of time. Just looking at initial costs without other factors that increase in costs as time progresses will be an incomplete analysis.

It is also prudent in the consideration is to include information from experiences in making this transition. Expanding usage of cloud storage has yielded useful information that can be applied as a general guideline. Here are some elements to consider when making a decision for on-premises and cloud storage.

Cost

The cost of storing data is pointed out first in almost any analysis. At Evaluator Group, we look at the practices around storing and managing information, which is more expansive than just the cost of storage alone. Examining costs factors requires looking at several potential areas.

- Opex vs Capex – Often cited, earlier determinations assumed that on-premises storage was a purchase-only option and compared that with per-gigabyte per-month charge for different classes of storage from cloud providers. Storage vendors have answered that challenge with consumption cost models where only the capacity used is charged as well as options for the vendor to manage the storage. The arguments for capex vs. opex for storage have been muted with many different options available.
- Orphan Data – Sometimes data is left on storage devices but no application “owns” the data. This is known as orphan data but also is called zombie data. In these cases, the identity of the data and its association for ownership is not clearly understandable. In on-premises environments, storage administrators manage allocation of capacity closely and regularly use tools to identify and deal with orphan data. A number of users of cloud storage have reported on orphan data where virtual machines or containers have been deleted but the data persists and is not easily identified. The reported numbers are in the range of 30% of storage capacity paid for is orphan data. The significance of this amount becomes pronounced in a TCO calculation spread over a number of years with increasing capacity demand.
- Multiple Copies of Data – Many of organizations that have moved applications and data to public clouds have also decided that administrators to manage the information were no longer necessary. After some analysis, they have reported the same applications and data that had been on-premises experienced a multiple of the amount of capacity in the public cloud. The multiple was reported as high as 5x and was attributed to information not being assigned to an administrator for management and the condition where

multiple copies were made for a variety of reasons. Having more data than necessary can inflate the costs for cloud storage significantly and the compounding effect will clearly show in a TCO projection.

- Physical Space and Environmentals – On-premises storage requires space and cooling which is a component of TCO. The fact that cloud storage charges include those costs is sometimes lost when doing a comparison. This again points to use of TCO as the most relevant comparison.
- Management of Physical Storage – On-premises storage systems and operations require management. The management includes updates required for storage systems and software used as well as transitions when a new system is installed and an older system retired. Those are direct costs either for the organization administrators or through a services group. For cloud storage, these functions are applied at a larger scale and expected to be significantly less. Any costs are applied in the charge per-gigabyte per-month for storage.
- Operational Changes – An unquantified cost difference between on-premises and cloud storage is when operational changes are made because of cost characteristics. The more obvious example is egress changes for some types of cloud storage. Because of additional charges, operational procedures may be changed to minimize charges. Any change represents some amount of time for operations to implement.
- Total Cost of Ownership (TCO) – As noted, TCO over a period of time is one of the best tools for understanding costs for storing and managing information. For analysis of costs of on-premises vs. cloud storage, the projection should be for seven years, which is the more modern expectation of lifespan for solid state (currently all flash) storage systems. That period of time, coupled with the expected capacity required, and the inefficiencies experienced will yield interesting information for evaluation.

Cost comparisons are only one aspect while considering on-premises or cloud storage. However, it is the most cited element in determinations.

Security

Security for information involves a number of different disciplines and crosses organizational boundaries. There have been significant advances with security capabilities for public clouds along with claims providers have made. The most important consideration for organizations/companies is that security is out of their control – the implementation and practices required for needed security lie with the cloud provider.

It is true that the cloud provider has staff with necessary expertise to address security demands. An organization/company can audit to the best available information provided regarding security but they must rely on assurances of the cloud provider that security practices are being maintained at all times.

Availability Assurance / Reliability

Obviously having information available for use by applications and the infrastructure reliable so that usage when required is not impeded is very basic. On-premises environments have staff

who work to make this the normal case and staff are available when an extraordinary situation arises. The cloud provider has staff to assure availability and reliability as well but experience has shown there have been significant impacts from different causes. The issue for many organizations is their helplessness in doing anything during an outage to regain business operations.

Data Protection

Data Protection is sometimes simply referred to as backup but is a more involved process to protect information assets which includes recovery capabilities. Because of threats to organizations through ransomware and other malicious acts, data protection is sometimes grouped with security but protection should be a normal business process and utilized when there has been some security incident. The security involvement should be to ensure that possible needs are being met through the data protection process.

Many of the initial usages of cloud storage was for independent groups with projects or applications. The assumption at that time was made that data protection provided by the cloud provider was adequate for their need. With IT organizations, the governance requirements for organizations overall and the security needs must be considered and determinations made as to whether the cloud provider offerings meet those requirements, are reliable, and can integrate into the corporate operational practices. This more detailed examination is required to make an informed decision regarding information asset protection.

Compliance / Privacy

Some information is under regulatory compliance regulations. It is the responsibility of the owner of the data to ensure that compliance requirements are met; not a task that can be transferred to another party such as a cloud storage provider. The cloud provider may provide storage that meets certain defined compliance requirements but the owner must constantly assure that is the case. Many of the requirements for regulations are about operational practices. This must also be taken into account when determining whether compliance requirements can be met.

Most companies are also concerned about privacy of data for many valid reasons. Related to compliance, assurances from cloud storage providers of maintaining privacy of data is important but the owner of the data must be responsible to verify that claim and provide the mechanisms to validate it.

Performance

Applications that do not seem to be performing to expectations require investigation and in the case of on-premises storage, this involves administrators with experience and skill to diagnose issues. In cloud storage environments, different tools are required because of the different environment. There needs to be staff with expertise to investigate. The considerations for dealing with perceived performance issues must also enter into plans and decisions.