Total Cost of Ownership (TCO) Model for Storage

Jonmichael Hands
SNIA CMSI SSD Special Interest Group Co-Chair
Sr. Strategic Planner, Intel Corporation

Table of Contents

Abstract

The total cost of ownership (TCO) provides a way to understand how customers make purchasing decisions by looking all the costs associated with purchasing and running IT equipment. Business objectives and purchasing criteria vary greatly between customers, and hence a one size fits all TCO model is not adequate for all storage workloads. We define a TCO model that is well suited for storage by looking at:

- Capital expenditures (CapEx) of the storage and everything else in typical datacenter rack that has typically dominated IT budgets, and
- Model operational expenditures (OpEx) which are dominated by power, cooling, and cost of device failures. This is useful for comparing a solution where the sole purpose is durable storage.

What Goes Into a TCO Model?

The total cost of ownership for data centers is broken down into capital expenditures (CapEx) and operational expenditures (OpEx).

CapEx is usually constrained within the data center and can include the price of hardware acquisition (compute, network, and storage), IT devices to manage, software and operating systems, as well as installation fees. The CapEx of the data center itself, which includes construction, real estate, government, taxes, fees, and business growth vary widely by a company. The current TCO model holds the datacenter CapEx to a constant input as Rack Cost per Year.

OpEx ranges from power consumption, depreciation of equipment, maintenance, repairs, IT software licensing, and data center staff costs.

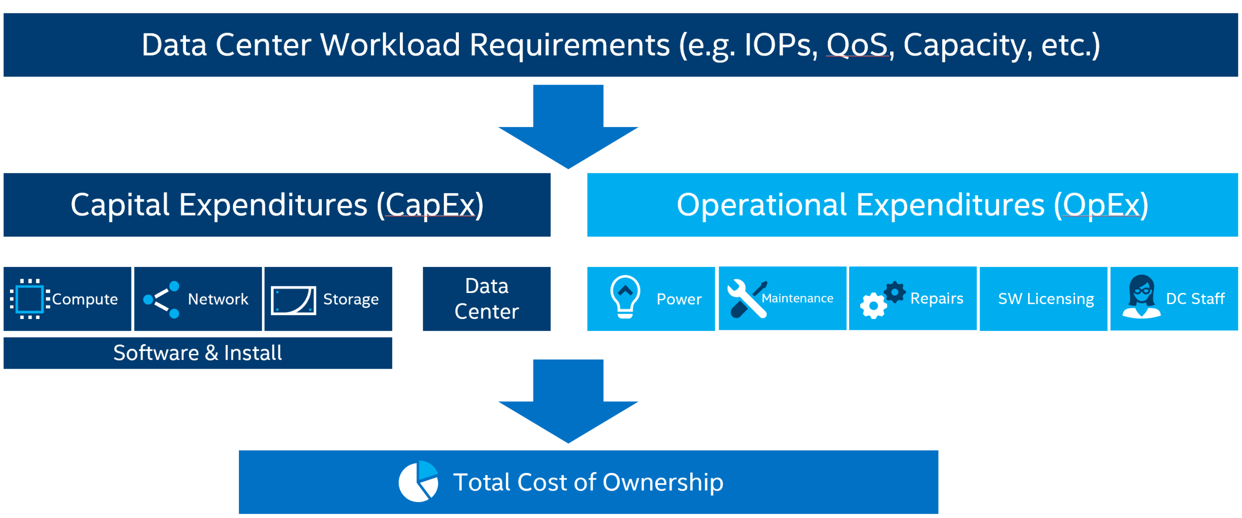

Figure 1. Total Cost of Ownership

Effective capacity is noted TBe, pronounced "terabytes effective". This is the actual usable storage space after replication, capacity utilization, and data reduction (compression, deduplication, etc.). This has a massive effect on TCO due to the multiplying of the entire CapEx of the raw storage. It is common practice for enterprise storage, all flash array, and others to advertise the effective capacity when looking at cost / TB of storage.

TCO Metrics

There are many metrics that may be important when modeling TCO for storage, including total system level performance, endurance, and total capacity. This model specifically looks at the TCO in TCO $ / TBe / rack / month. Large and small data centers alike use this as a reference to compare current vs future deployments that are optimized for the sole purpose of storing durable storage. TCO is comprised of the CapEx per TBe, which takes the total cost of all the IT hardware and rack and sums them with the total cost of storage, them amortizes that across the total amount of effective capacity. The OpEx is comprised of the power, cooling, and drive failures and is similarly amortized across the effective capacity per rack. This model does not include software licensing costs, but it could easily be added to any of the static fields in the rack configuration costs if required.

Modeling the Rack Configuration

There are many metrics that may be important when modeling TCO for storage, including total system level performance, endurance, and total capacity. This model specifically looks at the TCO in TCO $ / TBe / rack / month. Large and small data centers alike use this as a reference to compare current vs future deployments that are optimized for the sole purpose of storing durable storage. TCO is comprised of the CapEx per TBe, which takes the total cost of all the IT hardware and rack and sums them with the total cost of storage, them amortizes that across the total amount of effective capacity. The OpEx is comprised of the power, cooling, and drive failures and is similarly amortized across the effective capacity per rack. This model does not include software licensing costs, but it could easily be added to any of the static fields in the rack configuration costs if required.

Drive Input

The model is split up into hard disk drive (HDD) drive, and solid state drive (SSD) inputs. These correspond to the rack configurations for each setup. Input is required for the capacity of the drive, active and idle power specification (which are used with the duty cycle to estimate total power consumption), annual failure rate in % (model drive failures and replacement costs), and cost in average selling price (ASP) and expressed in $/GB (which need to match).

Workload Input

Drive Performance Impact on TCO

SSDs are of course much faster than HDDs, so looking at an ***application performance*** driven TCO model will most likely favor SSDs when trying to achieve a specific SLA or number of IOPS. Workload inputs to the TCO model that materially change the output are capacity utilization factor, error encoding factor, multi-workload replication factor, and data reduction technology.

Performance multiplier is the benefit of SSDs over HDDs in performance or QoS represented as a multiplier of HDDs that are needed to hit a certain IOPS, QoS, or bandwidth and is an easier way to express application performance differences between devices. When data is more frequently accessed the amount of IOPS/TB required increases. Replication schema (covered below) may also impact performance, so be careful not to double dip since this is a multiplier in the model.

Capacity Utilization is the percentage of a total storage device that is used out of the total available capacity. Generally, drives are left with some free space due to the performance impact of being full and to ensure ample free capacity for incoming data.

Error encoding / replication takes into account the replication and data protection schema like RAID or erasure coding. For example, mirroring data would yield a factor of 2, triple replication would be 3. SSDs often have an advantage over HDDs because of reliability, rebuild performance and durability that allows thinner overhead with equivalent or better data durability. A reference tab has been included for common overheads of RAID 5, RAID 6, and erasure codes.

Data Reduction Ratio is modeled as percentage reduction and is the ratio of host data stored to physical storage required. E.g. a 50% ratio would be equivalent to a 2:1 data reduction ration. Since data reduction allows the user to store more data then is on the physical hardware, the resulting effective capacity is increased. Technologies such as compression, deduplication, can greatly decrease the required raw capacity needed to meet a “usable capacity” requirement. SSDs have an advantage vs HDDs in compression with modern algorithms and offloads due to the higher performance (IOPS and bandwidth). This is highly dependent on performance and frequency of access of the data, and where the compression is taking place (outside of the scope of this model). While compression ratios are algorithm dependent, not storage device type dependent, the compress and decompress speeds will be limited by disk access speed (IOPS). For industry leaders in all-flash arrays, data reduction is their value-add in customized software and they often cite all-flash being what enables high levels of DRR while delivering high performance. Another example is VMware vSAN, where compression and deduplication are only offered in all-flash configurations. Modern compression algorithms, for instance, ZStandard from Facebook, can achieve compress and decompress speeds much faster than HDDs can read/write thus allowing the use of the algorithms on SSDs in real time. It also offers benefits for flash with small file compressibility through dictionary files to significantly improve database compression performance. A common use of PCIe accelerators, FPGAs, offloads, and Intel® Quick Assist Technology for doing real-time encryption, compression, and deduplication provide a tremendous benefit with high-performance NVMe SSDs due to the real-time requirements and high throughput required.

Duty Cycle is the percentage of the time the drive is active over the deployment. This is solely used for active and idle power assumptions. In hot or warm storage that is frequently accessed, latency is important and devices never spin down or enter near zero idle power states. In colder or archive storage that is much less frequently accessed lower idle power settings may be used.

Operational Expenditures

Power consumption is one of the largest drives of operational expenses in a data center. Power is also a major constraint about how a rack is designed and managed. OpEx from power in the model is calculated by summing the power consumption of IT equipment in the rack (servers, switches, JBODs, etc.) and the total power consumption of the storage. The power consumption of the storage is modeled with a duty cycle that dictates how often the devices is in the idle state vs active state, and estimates total amount of energy consumed in the deployment term. After the total amount of power consumed is calculated, the energy consumption in kWh and cost are calculated with fixed constants in $/kWh of electricity and data center PUE. PUE is defined as the ratio of the data center total energy consumption to information technology equipment energy consumption, calculated, measured or assessed across the same period[1] . In less efficient data centers a larger PUE is typical, as compared to hyperscale data centers that have a very low PUE and typically have a larger amount of energy consumed from renewable energy.[2]

Cooling costs – a typical data center spend a large amount of OpEx on cooling costs. This model just assumes the cooling costs are built into the IT equipment (fans, liquid cooling, etc.) rather than looking at it from a data center level. If the cooling cost is known and fixed for a given rack power, it can be added in the rack section to the fixed data center costs to be amortized over the given amount of storage capacity.

Drive failures are inevitable when deploying a large amount storage. Device manufacturers typically list their failure rate in MTBF (mean time between failure) or AFR (annual failure rate). These are mathematically related expressed by MTBF = 1/AFR*365*24. SSDs are known to have lower failures than HDDs due to the fact that they have no moving parts and are less prone to mechanical wear and tear.[3] SSD failures are well understood and have quite different failure mechanisms from HDDs.[4] [5] HDD failures often exceed what is listed by the vendor, with independent studies showing an average of 1.58% AFR over the previous 7 years of deployment.[6] In non-ideal conditions, at high temperature, humidity, or exceeding the HDD rated workload (similar to endurance on SSDs) HDD failures approached 4% AFR.[7]

The model calculates the total number of drives from the rack input, and finds the total number of drive hours my multiplying by the deployment term. This is then divided by the MTBF in hours to estimate the number of device failures in the deployment term. A fixed replacement cost is given to send a technician out to service the failed device. Enterprise HDD and SSD are assumed to be under a 5 year warranty thus no drive replacement costs for the actual storage is assumed.

---

[1] ISO/IEC 30134-2:2016. Information technology — Data centres — Key performance indicators — Part 2: Power usage effectiveness (PUE)

[2] https://tech.fb.com/hyperefficient-data-centers/

[3] Mielke, N., Frickey, R., Kalastirsky, I., Quan, M., Ustinov, D. and Vasudevan, V., 2017. Reliability of Solid-State Drives Based on NAND Flash Memory. Proceedings of the IEEE, 105(9), pp.1725-1750.

[4] Reliability of NAND-Based SSDs: What Field Studies Tell Us

[5] A Study of SSD Reliability in Large Scale Enterprise Storage Deployments

[6] https://www.backblaze.com/b2/hard-drive-test-data.html

[7] Feeding the Pelican: using archival hard drives for cold storage racks

About the SNIA

The Storage Networking Industry Association (SNIA) is a non-profit organization made up of member companies spanning information technology. A globally recognized and trusted authority, SNIA's mission is to lead the storage industry in developing and promoting vendor-neutral architectures, standards and educational services that facilitate the efficient management, movement, and security of information. More information is at www.snia.org.

About the Compute, Memory, and Storage Initiative

The Compute, Memory, and Storage Initiative supports the acceptance and growth of computational storage (CS) solid state storage (SSS), and persistent memory (PM) in the marketplace. Our member companies support SNIA work in their technology focus areas of computational storage, physical storage, and persistent memory; promote their results in the marketplace, influence standards activities, and educate vendor and user communities. Find out more at www.snia.org/forums/cmsi.

Storage Networking Industry Association
4360 Arrows West Drive • Colorado Springs, CO 80907 • Phone: 719-694-1380 • Fax: 719-694-1389 • www.snia.org