Enterprise Storage: “Hot” and “Cold” Data

What is Data Temperature?

Data and its storage needs can be simplified across a temperature spectrum with the extremes being “hot” and “cold.” These two sides refer to the characteristics of the data and its requirements. Data is “hotter” if it is most-accessed, mission-critical, or has stricter performance requirements. Typically, this is often-accessed data that doesn’t have time to “cool” and therefore must be readily available. Data is “colder” if it is not as frequently accessed, if the data is less important, or if extraction speed is less

critical.

In-between the hot and cold metrics there's the possibility for additional categories, for example, defined as "warm" and "cool." This allows further delineation of data for flexible organization. This can make sense as colder data tends to require more capacity but can manage with lower performance, for example. Typically, the choice in designation is based on the age of data – with older usually being colder – but also workload type, with oft-churned data being hotter. Segregating data along these lines can make for a more efficient and cost-effective storage scheme.

Storage Tiering

With a potentially diverse level of requirements for storage across this temperature spectrum, the natural inclination is to build a hierarchical or tier-based system. The hotter tiers require faster access speed and lower latency, higher input/output operations per second (IOPS), and may contain higher-value data. This calls for faster and more manageable storage as you get with non-volatile media express (NVMe™) solid state drives (SSDs). Faster access, as with on-premises drives versus a non-hybrid cloud, may also be a requirement. These drives will likely be more expensive due to these necessities, including power loss protection (PLP).

Many storage hierarchies in fact include tiers with dynamic random access memory (DRAM) or persistent storage. In both cases, capacity is limited and tiers with different types of memory can add complexity. In either case, you still need a colder tier with lower performance requirements. The replacement of HDDs for SSDs continues unabated here, particularly with serial ATA (SATA) SSDs that offer capacity at a good price versus more expensive SSDs while maintaining a lower total cost of ownership (TCO) than HDDs. As access is less critical and can be slower, this includes the use of cloud services including hybrid cloud.

Software and Optimization

Aside from hardware, optimization and automation for hot and cold data can be achieved through the use of software and design. Data must be segregated in an intelligent and efficient fashion so that some data does not end up in the wrong tier. Many organizations may require focused data delineation, relying on finer-grain control, while others can simplify by looking only at data age or workload intent. The correct solution depends on the overall storage system plus the type of data and operations utilized on the data. In this respect, choosing the right SSD requires a cost-benefit analysis with the knowledge that software can improve efficiency by utilizing storage to its maximum effect, if necessary.

This can include information as derived from the file system (FS) or host, for example. The NVMe™ specification contains the streams directive which can direct writes to intended zones based on metadata; see our white paper for more information. This can be combined with the pre-sorting of data, including by the application or applications utilizing the data. Data location – over the network, in the cloud, et cetera – and data abstraction type, for example with virtual machines (VMs), can also influence this process. Those not requiring fine-grained control can therefore rely on SATA SSDs instead.

Summary

Enterprise data can be divided into temperature zones based on the immediacy requirements of the data. This goes from “hot” to “cold” with possible intermediate states such as “warm” and “cool.” Generally, hotter data is accessed more frequently with correspondingly higher performance requirements. Colder data may contain older data and capacity often becomes more of a concern. Tiering can include other types of memory, such as DRAM or even HDDs, but generally, with non-volatile storage, we’re distinguishing between NVMe™ and SATA SSDs within the storage ecosystem.

Organizations can control the placement of data in multiple ways, from simple to complex, utilizing everything from NVMe™ directives to application-level control, and at various levels of granularity. This can be dependent on the overall storage system as well as the data and workload type. Efficient and

effective segregation of the data is crucial for making the most of the storage – to save on cost by using the correct SSDs for the job. Whether the data is kept on-site, in the cloud, or both, the choice of storage is important to find the right balance between cost and data availability. The first step is to determine your storage priorities, find out the performance, capacity, and access needs, then dial in to the right SSDs from our portfolio.

Our SSD Solutions

PCIe™ - Our ED1 Series is a powerful, high performance SSD made for edge storage applications. It comes in M.2 and U.2 form factors.

SATA - Our ER2 SSD Series delivers affordability and performance with superior random read/write speeds of up to 90,000/45,000 IOPS. It comes in M.2 and 2.5" form factors.

Please contact our Solid State Storage Technology Corp. expert for more information.

*Specifications and features are subject to change without prior notice. Images are samples only, not actual products.

Request Full Specs Sheets

ABOUT US

A subsidiary of KIOXIA Corporation, **Solid State Storage Technology Corporation** is a global leader in the design, development, and manufacturing of digital storage solutions. We offer a comprehensive lineup of high-performance customizable SSDs for the Enterprise, Industrial, and Business Client markets. With various form factors and interfaces, our SSD solutions help businesses simplify their storage infrastructures accelerating variable workloads, improving efficiency, and reducing total cost of ownership.

Learn more at www.ssstc.com