2021 Storage/Computing Trends

SSDs: Continued Growth in 2021 and Beyond

Artificial Intelligence: Driving Future Storage Architectures

Will 2020’s Hot Topics in Storage Still be Hot in 2021?

Cheers to the New Year:

Maybe I am stating the obvious, but I don’t think it is often said – all of you make the visions and dreams of innovators like Steve Jobs, Bill Gates, Elon Musk, Jeff Bezos, Reed Hastings, Mark Zuckerberg possible. The fact is that every big idea of today involves the storage, movement, management and availability of huge amounts of data quickly, efficiently, and decisively. The amount of data keeps growing, the applications of this data seem endless, and the speed and reliability of accessing this data has become critical in terms of economics, need, and in many cases such as space exploration and military applications, life-saving.

2020 has been a unique year, and companies and industries have been quick to evolve to meet the demands of remote work, while keeping the economy moving and protecting the safety of employees and customers. We had our first virtual conferences, exploring new methods to network and market products. Webinars became even more integral to those efforts.

This newsletter looks at major areas of enterprise storage and the trends for the 2021. For example, in the area of storage management for healthcare, hospitals are dealing with vast amounts of data from many different sources, in varying formats (physician notes versus an MRI, for example), security challenges, and AI-modeling for diagnosis. Add to those complexities, innovation in medical treatment and equipment, and, even a pandemic, and enterprise storage plays a pivotal role.

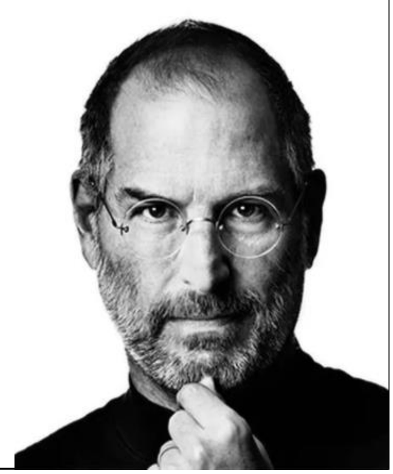
Everyone truly appreciates the sacrifices of medical essential workers, staffing hospitals each and every day, while much of the world works safely from home. It is challenging to adequately express our gratitude. But, I would be remiss if I didn't also acknowledge that it is your contribution to data management that provides quick turn around for worldwide tracking of outbreaks, data availability for hospital treatments and care, and for the data analytics necessary in creating vaccines. Effective data management touches all facets of our daily lives.

Thank you for your contributions to the enterprise storage industry and making the data-driven dreams of innovators a reality.

Cheers to a successful and healthy 2021!

Mike Heumann and the entire G2M Team

"I love that the work that we do is so vital to science. We're in a lot of ways at the scientific front line. The work that we're doing to build up the computational defense system for infectious diseases, whether it's finding the vaccine as fast as possible this time or next time to detect early outbreaks." Jensen Huang, CEO, NVIDIA

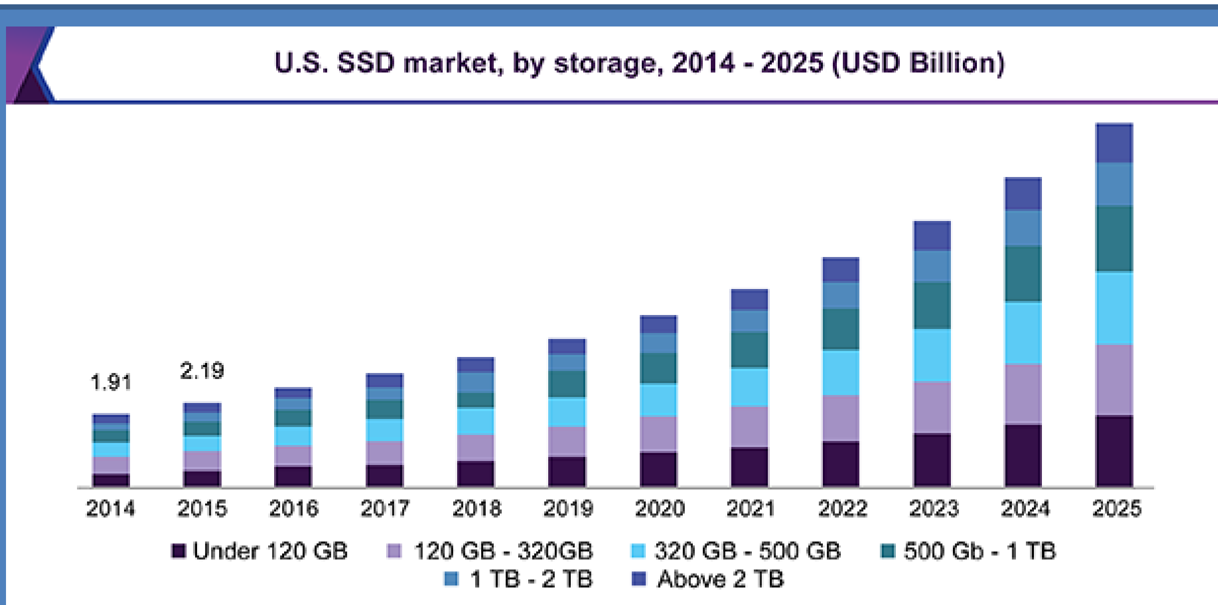


"Your work is going to fill a large part of your life, and the only way to be truly satisfied is to do what you believe is great work." Steve Jobs, CEO, Apple.

SSDs: Continued Growth in 2021 and Beyond

Clearly, the future of enterprise storage is solid state drives (SSDs). In 2020, the dominant leaders in enterprise SSDs were (in order of market share) Samsung, Intel, SK Hynix, Western Digital, Kioxia, and Micron. SSDs are now the primary storage media for mainline storage, and are making significant inroads into nearline storage as well. SSDs are clearly continuing to increase in capacity, doubling about every 12-18 months. The problem that this capacity increase poses for enterprise storage architects is that most server use cases do not require the capacity levels that SSDs are now reaching; sharing direct-attach SSDs across multiple servers is problematic.

It will be interesting to see whether 1) network-attached SSDs (basically mini-storage arrays) take off; 2) scale-out flash storage (SOFS) software continues to grow and becomes more the norm for sharing; and/or 3) manufacturers use form factors like EDSF to build physically smaller SSDs, but with lower capacity. The logical direction is the second option given the flexibility that software approaches provide, but companies have generally been resistant to integrating hardware and software themselves. One option is for the larger storage and server companies to purchase SOFS companies and then sell the software in their storage products. Alternatively, many of the storage/server companies (and SSD companies, for that matter) may choose to enter SOFS software market themselves, as Kioxia has done with its Kumoscale software.



https://www.grandviewresearch.com/industry-analysis/solid-state-drives-industry

Artificial Intelligence: Driving Future Storage Architectures

We have seen how the use of Artificial Intelligence (AI) and Machine Learning (ML) are driving the need for both storage performance and capacity. Applications such as medical radiometry, facial recognition, autonomous vehicles (AVs), e-commerce, and machine vision require extensive datasets for training. As AI/ML solutions are “iterated” on to improve performance and/or adjust for changing circumstances, it is also important to retain the previous training and verification datasets, further increasing the amount of storage required. And, while AI and ML models generally execute quickly, the construction of those models (“training”) can take a significant amount of time. Since training times are proportional to the size of datasets and the speed of the storage that holds the data, often times the only way of reducing training times is to improve the performance of the storage holding the data.

Which brings us to the other way that AI and ML are driving future storage architectures. One of the hottest segments of application performance management (APM) is in the use of AI to analyze infrastructure performance. This is especially true in virtualized environments where failover, data migration, and similar activities can impact overall system performance. Storage infrastructure is no different, which is why enterprise storage solution companies have started utilizing AI to increase the performance and availability of their storage platforms. Initially, AI/ML was utilized for advanced diagnostics and provisioning – what is going to fail next, and where will I need to add more storage to the system? This is especially important for flash-based systems, where drive wear increases with performance demands (more writes, more wear).

Storage vendors are now going beyond these first-generation uses of AI and incorporating dedicated AI software and/or hardware to speed up specific applications. By monitoring storage performance simultaneously with infrastructure changes and the performance of specific applications, AI can identify what factors are impacting the applications. As mixed storage models (public cloud, hybrid cloud, and private cloud) become more prevalent, the need to measure the impact of “invisible’ cloud infrastructure changes on performance will be more critical. Expect to see cloud storage and infrastructure providers start to offer AI/ML based performance monitoring in the future.

"I take the position that I'm always to some degree wrong, and the aspiration is to be less wrong." Elon Musk, CEO, Tesla

"Stone Age, Bronze Age, and Iron Age. We define entire epics of humanity by the technology they use." Reed Hastings, CEO, Netflix

Will 2020's Hot Topics in Storage Still be Hot in 2021?

Here is a look at three topics for 2021 - storage-class memory, computational storage, and composable infrastructure - what they promised to deliver, to what extent they have lived up to their promise, and what to expect in the coming decade.

Storage-Class Memory (SCM): While the access speed of non-volatile storage as increase by orders of magnitude with the migration of hard disk drives (HDDs) to flash-based SSDs, the speed of DRAM has also increased, as has its price. This has resulted in what some industry pundits call the "RAM gap" - the price and performance discrepancy between DRAM and flash-based SSDs. Enter SCM, which promises to fill this gap with a technology that is 10X faster than SSDs, is non-volatile, and costs less than DRAM. While several SCM technologies have been fielded (including Intel's® Optane™, vSMP MemoryONE, mixed DRAM-flash devices, Samsung's Z-SSD, and others) none of them have taken off significantly, primarily due to difficulties with their programming model.

The one place that these technologies have taken off is in storage arrays, where they can significantly reduce hardware costs by reducing the RAM needed. There are new technologies on the way (MRAM, NRAM, PCM, and STT-RAM) that promise to provide non-volatile byte-addressible storage that might meet the need that SCM has always promised to address.

Computational Storage: One of the biggest delays in analyzing extremely large data sets is the movement of data from storage such as SSDs to DRAM in the server. In the case of transfers from local SSDs, this speed is limited by the speed of the PCI Express® (PCIe) bus (3.94GB/s for the PCIe Gen3 x4 that is typical for U.2 SSDs). To put this in perspective, it would take about 26 minutes to transfer 100TB of data from sixteen (16) PCIe Gen3 U.2 SSDs. Computational storage turns this problem on its head by saying "what if you don't transfer the data, but actually perform the processing in the SSD?"

Of course, there are tradeoffs in this approach, chief among them being that you cannot put an Intel X86 processor inside of a U.2 SSD. Most SSDs utilize ARM processors internally (in some cases augmented by an FPGA), which means changes to the applications that will analyze the data. While this may be doable for third-party software, is is not likely for popular 'big data" programs like SAP HANA, limiting the applicability of computational storage. The ARM processors also do not necessarily have processing power on par with X86 processors in "big iron" servers, so there is definitely a tradeoff between reducing transfer speeds and processing speeds. With 400GbE networks coming out (48GB/s – 12 times the speed of PCIe Gen3x4) and the advent of PCIe Gen4 and Gen5 (especially with 8 lanes), the transfer time issue may be going away.

Where computational storage has found some legs is with embedded applications, where the SSD is more like very small footprint, low-power, hardened single-board computer. This is an important capability for a variety of use cases in aerospace, military, automotive, and industrial applications, where space and power are at a premium and an "encapsulated" system that is field-pluggable is highly desirable.

Composable Infrastructure: Composable infrastructure traces its legacy back to the previous decade (the "naughts"), where it was called "I/O virtualization". Composable infrastructure essentially disaggregates the entire server and allows the parts to be shared across programmable high-speed fabrics, including any of the flavors of NVMe-oF (including 100GbE), or a "fabricized" version of PCIe running NVMe. The fabrics incorporate a switching infrastructure that allows components to be connected and reconnected at will to various servers – essentially 'sharing' those components, reducing costs. In general, composable infrastructure makes the most sense when the components to be shared are both relatively expensive and sparsely utilized – things like GPGPUs or FPGA cards. Of course, this flexibility isn't free – running connections through fabrics always adds latency, and the increased amount of cabling required does add complexity to racks. Nevertheless, there are a good number of use cases that fit these "best case" parameters – media and entertainment and oil and gas modeling to name a few. If combined with smaller modular servers (which we will see more of in the coming decade), composable infrastructure can make a lot of sense.

"As we look ahead into the future, leaders will be those who empower others." Bill Gates, Microsoft

"What is dangerous is not to evolve."
Jeff Bezos, CEO, Amazon

G2M Research Multi-Vendor Webinar Series

Our 2021 webinar schedule is ready! Click on any of the topics to get more information about that specific webinar. Interested in Sponsoring a webinar? Contact **G2M** for a prospectus.

Our November webinar Implementing NVME™ & NVMe-oF™ for Cloud Service Providers was sponsored by Kioxia (Joel Dedrick), Lightbits (Josh Goldenhar), and Western Digital (Mark Miquelon). View the recording and/or download a PDF of the slides.

2021 Webinars

Jan 19: Can Your Server Handle The Size of Your SSDs?

Feb 23: Storage Architectures to Maximize the Performance of HPC Clusters

March 23: One Year after COVID-19: How Did Storage Architectures Perform for Biotech AI Modeling & What Can We Learn From This?

April 20: The Race to be Relevant in Autonomous Vehicle Data Storage (both On-Vehicle and Off-Vehicle)

May 18: Responsive and Efficient Storage Architectures for Social Media

June 15: It's 2021 - Where Has NVMe-oF™ Progressed To?

July 13: Computational Storage vs Virtualized Computation/Storage in the Datacenter: "And The Winner Is"?

Aug 17: AI/ML Storage - Distributed vs Centralized Architectures

Sept 14: Composable Infrastructure vs Hyper-Converged Infrastructure for Business Intelligence

Oct 12: Cloud Service Providers: Is Public Cloud, Private Datacenter, or a Hybrid Model Right for You?

Nov 9: The Radiometry Data Explosion: Can Storage Keep Pace?

Dec 14: 2021 Enterprise Storage Wrap-up Panel Discussion