# VMware VMworld 2021 announcements

**Silverton Consulting, Inc. StorInt™ Briefing**

At VMworld 2021, VMware announced a many updates including VMware vSphere 7 Update 3, VMware vSAN 7 Update 3 and additional projects to be rolled out over the coming months/years. Below we discuss 3 announcements we believe most relevant to the enterprise today.

## vSphere 7 Update 3 support for NVMeoF/TCP

vSphere has had support for NVMeoF/FC and NVMeoF/RoCE for years now. But RoCE has always required special, more expensive RDMA compatible switches and HBAs, which for greenfield data center are not much of a problem but for existing data centers, there's a significant upgrade cost involved.

Enter NVMeoF/TCP, a technology we first discussed back in 2018 with SolarFlare on a GreyBeards on Storage podcast. Back then it was still an immature technology but it's now ready for wider adoption. Probably the most interesting difference between NVMeoF/RoCE and NVMeoF/TCP is that it can be used over current HBAs and Ethernet switches without modification.

Early on there was some discussion that NVMeoF/TCP might be slightly slower than RoCE storage. While this may be true even today, the differences are probably unlikely to matter to enterprise users. If NVMeoF/TCP adds 5-10% latency to an IO operation vs. NVMeoF/RoCE, yes, that's significant but just barely. And when we are talking about an IO operation which today might take 300 µsecs or more on SAS/SATA storage, which on NVMeoF/TCP might take 110 µsecs, you are still getting a lot of performance boost for essentially a software upgrade and new SSDs. Of course, with an Optane SSD, you might be able to see an 11-12 µsec response time with NVMeoF/TCP.

## vSAN 7 Update 3 support for HA k8s storage

With vSAN 7 Update 3, there's three new features that may be of interest to K8s users.

- **vSAN with Tanzu with vSphere support for file based, RWX (read-write-many) PVs** (persistent volumes). This means the same file RWX PV can be accessed and updated from multiple k8s pods. vSAN with Tanzu already supports RWO (Read -Write Once) block-based PV.
- **vSAN supports stretched clusters with K8s**. k8s containers can now access persistent volumes through vSAN across stretched clusters. In the event vSAN storage goes down on one cluster, it's replicated cousin can be accessed on the other cluster. This includes **site affinity** support for k8s persistent volumes as well as support both RAID-1 and RAID-5/6 data protection at secondary sites.
- **vSAN supports K8s topology aware volume provisioning**. In a multi-cluster K8s and vSAN environment, PVs can be automatically provisioned in vSAN clusters that match their K8s cluster topology. Topology aware PV provisioning should make supporting multi-cluster K8s environments easier. This support is available through the CSI provider and only available in vanilla K8s environments.

## Project Capitola for memory tiering virtualization

Lately, server memory has become a lot more complicated, what with Intel PMEM, CxL (Compute Express Link) memory and SmartNICs/Sophisticated GPUs. Customers applications need access to very large memory spaces, which are expensive and can only be accessed now in direct mode, isolated to a single server and VM.

PMEM and CxL, memory need no longer be a single server/VM resource anymore. CxL currently operates over a PCIe Gen 5 bus which enables a shareable, hot pluggable memory and PMEM could easily be shared across all the VMs in a ESXi host. But In order to share PMEM/CxL memory across VMs, there needs to be a virtualization layer between the advanced memory hardware and the VM. That's where VMware's **Project Capitola** comes in, it essentially provides a virtualized, software defined memory tier, that includes DRAM, PMEM and (CxL based) NVMe memories for VMs.

With Project Capitola functionality, VMs can allocate and provision memory on the fly. The fact that some memory might reside in the ESXi server while other memory may be out on a NVMe attached device shouldn't matter to the application anymore. Project Capitola will roll out first with memory tiering for VMs in the same ESXi host and later with similar capabilities, that spans the whole vSphere cluster.

## Significance

There's a spate of storage vendors announcing NVMeoF/TCP iSCSI support including but not limited to Dell PowerStore and NetApp ONTAP storage systems, but LightBits Labs has had this since coming out of stealth. But what's impressive is that for most storage vendors, NVMeoF/TCP is only a software change. As such, with vSphere 7 Update 3, another software change, customers can now take advantage of much faster response times and higher IOPs without hardware changes. Nutanix and Red Hat OpenShift seem to not be there yet, but it shouldn't take them long to catch up.

vSAN's continuing roll out of K8s capabilities is interesting. VMware, early on, realized that support for K8s enterprise applications would require stateful containers. The race has always been between storage vendor's and vSAN's K8s support. vSAN's support for RWX file PV's will go a long way to making it more of an even race between storage vendors and vSAN.

Customers are demanding to run ever more memory intensive applications, which can all be done today, on ESXi with its support for Intel Optane PMEM in an isolated fashion. But PMEM's not going to be the last word on memory technology. In order to fend the proliferation of new memory technologies and make that hardware be accessible to all VMs in a host, something like Project Capitola memory tiering virtualization needs to be in place. Getting there soon is a necessary evil, when VMware's competing with the public cloud, native K8s and other virtualization products.