Technical Validation

Deploying Multi-fabric Architectures with Liqid Composable Disaggregated Infrastructure

Configuring Servers On-demand in the On-premises Data Center

By Alex Arcilla, Senior Validation Analyst

June 2022

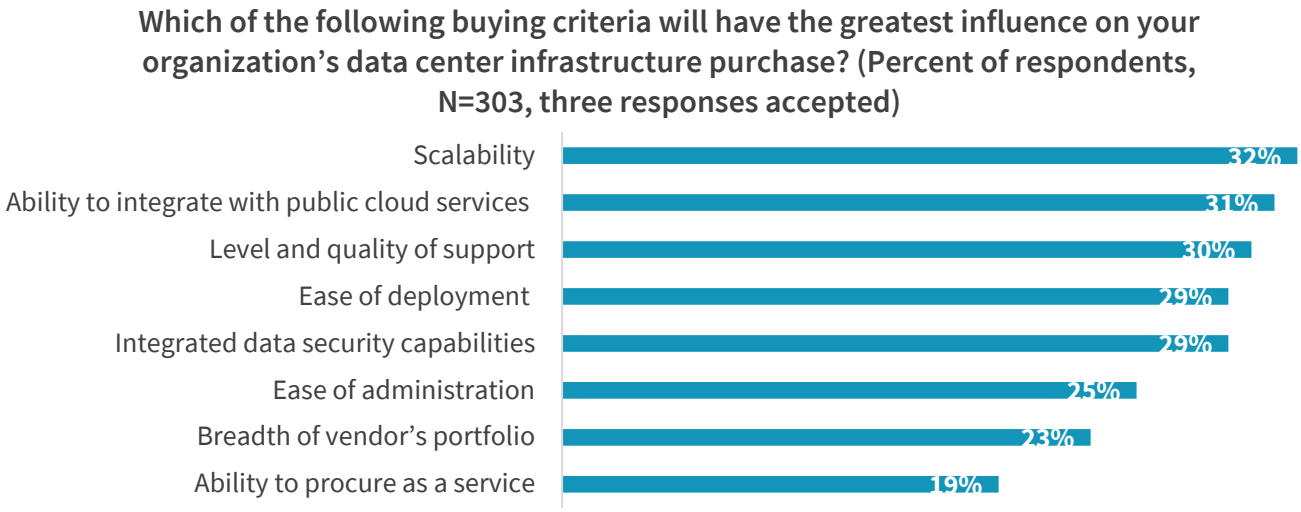
Introduction

This ESG Validation Report documents testing and evaluation of Liqid Composable Disaggregated Infrastructure (CDI), with a goal of validating how organizations can accelerate time to value, increase resource efficiency, and improve IT agility when deploying IT resources, including NVMe over Fabrics (NVMe-oF)-based storage architectures, within their on-premises data centers.

Background

Organizations are under pressure to move as quickly as possible to meet business needs, as ESG research uncovered that 66% of respondents believe that they have to move between 25% and 100% faster to deploy applications, infrastructure, and services. This pressure has definitely influenced organizations' buying criteria for data center infrastructure, as ESG research found that 29% of respondents consider the ease of deployment as one of the top criteria that will have the greatest influence on their data center infrastructure purchases (see Figure 1).[1]

Figure 1. Top 8 Criteria Influencing Data Center Infrastructure Purchase



Source: ESG, a division of TechTarget, Inc.

Deploying data center infrastructure has been a significant bottleneck since the process of purchasing, installing, and configuring equipment takes weeks to months. Moreover, re-deploying that same equipment is not straightforward as workload requirements vary. As a matter of fact, ESG research found that 63% of respondents still find it difficult to properly size workloads for the optimal infrastructure environment, leading to either over- or under-provisioning.

With infrastructure-as-a-service (IaaS), organizations have been able to deploy and scale virtualized servers on-demand in the public cloud to gain advantages of speed, agility, and scalability. Yet, sizing workloads correctly remains an issue when using cloud compute instance types with predefined amounts of CPU or GPU capacity, RAM, and storage. The risk of over- or under-provisioning sufficient instances for a workload persists, directly impacting both workload performance and overall expenses.

To meet business requirements quicky without increasing expenses unnecessarily, what if organizations could configure and scale physical servers on-demand in minutes, bypassing the traditional IT provisioning and installation process? When

[1] Source: ESG Research Report, *2021 Data Infrastructure Trends*, November 2021. All ESG research references and charts in this ESG Technical Validation have been taken from this survey results set, unless otherwise noted.

a workload is no longer needed or requirements change, what if organizations could reclaim unused resources and redeploy them to support new workloads?

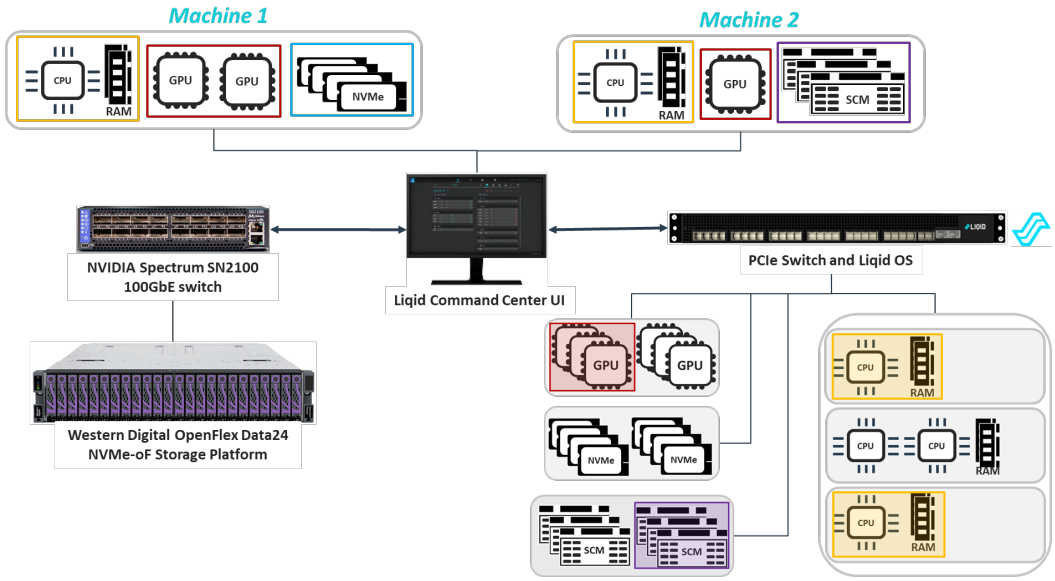
Liqid Composable Disaggregated Infrastructure

Liqid Composable Disaggregated Infrastructure (CDI) is a software-based solution that enables organizations to provision bare-metal servers on demand, without the burden of traditional IT provisioning and installation processes. The solution brings the flexibility and agility of IaaS to on-premises data centers by enabling organizations to quickly configure and modify servers from pools of hardware resources, including compute,[2] GPUs (e.g., NVIDIA A100), NVMe SSDs, storage-class memory (SCM),[3] FPGAs, and NICs, over PCIe fabric or Ethernet, in real time. These software-defined server configurations function as bare-metal servers that support commonly used OSs, hypervisors, or container engines.

With Liqid CDI, organizations can configure a server with compute power that would typically require multiple chassis, such as a single 1U server with 16 GPUs. To meet evolving workload requirements, they can quickly add or remove accelerator, storage, and/or networking resources via software, eliminating the need to physically install or remove components from a server chassis. CDI allows them to deploy and scale the exact quantity and type of resources that new and scaling workloads require.

Figure 2 illustrates how Liqid composes resources for a host server using both PCIe Gen 4.0 and Ethernet fabrics. This multi-fabric approach provides organizations with flexible connectivity options to compose servers with the required resources for a wide variety of workloads. To compose servers via PCIe fabric, a Liqid PCIe HBA is installed in each server and then they are connected to a Liqid PCIe switch via mini-SAS cables. When using Ethernet fabrics, they must support direct data access between systems to eliminate expensive memory copies via Remote Direct Memory Access (RDMA), thus maximizing NVMe storage performance.

Figure 2. Liqid Composable Disaggregated Infrastructure – Multi-fabric Architecture



Source: ESG, a division of TechTarget, Inc.

[2] CPUs and RAM are not disaggregated in Liqid CDI. Servers already contain a defined amount of CPU and RAM capacity.
[3] With Intel Optane technology, SCM components can extend system memory.

All disaggregated resources—GPUs, storage-class memory (SCM), FPGAs, and NVMe SSDs—reside in expansion chassis and are connected to their respective fabrics. For organizations that want to implement NVMe-oF storage architectures, Liqid partners with Western Digital to disaggregate OpenFlex Data24 NVMe-oF Storage Platform.

Residing on the switch fabric, Liqid Matrix software enables the allocation of resources to servers via UI, API, or CLI. Organizations can also use SLURM or VMware vCenter to compose servers. No Liqid-specific drivers or agents are needed for the server to recognize composed devices. Resources can be added to running systems without a reboot.

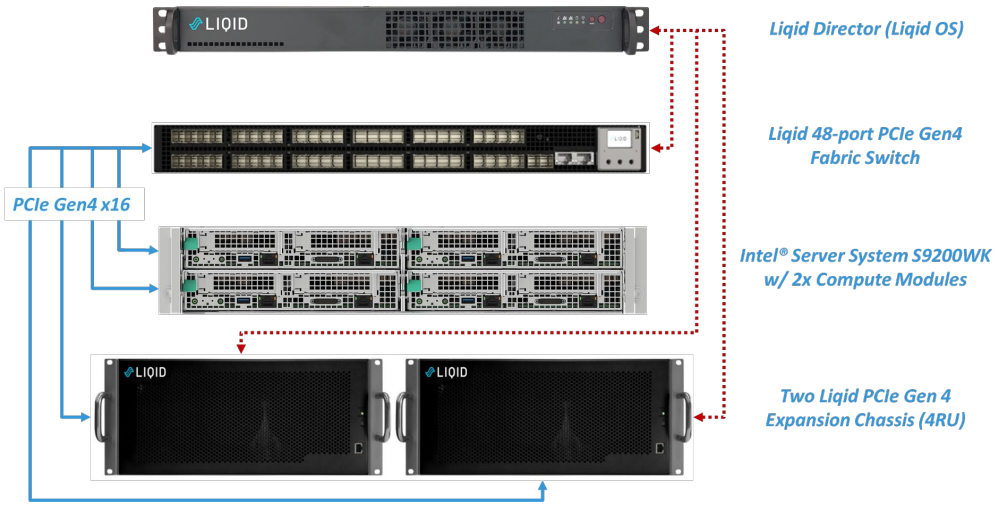
All components can be scaled independently of each other to meet any workload's requirements. If the workload is no longer running or requires fewer resources, those components can then be returned to the free pool for reallocation to other servers.

ESG Technical Validation

ESG performed evaluation and testing of Liqid CDI at the company's facilities in Broomfield, CO. Testing was designed to demonstrate how the solution can help organizations accelerate time to value, improve IT agility, and increase resource efficiency.

ESG began with the test bed shown in Figure 3. One Liqid 48-port PCIe Gen4 Fabric Switch interconnected with two compute modules within an Intel server system S9200WK equipped with Intel Xeon Processors. Each contained a Liqid PCIe Gen4 x16 Fabric HBA card. The Liqid Fabric Switch also connected to two Liqid expansion chassis, each containing eight PCIe Gen4-enabled slots. Both chassis were populated with a total of 16 NVIDIA A100 Tensor Core GPUs. The Liqid Director interconnected the Liqid Fabric Switch and expansion chassis to enable server composability.

Figure 3. Liqid Test Bed for Composing a New Server



Source: ESG, a division of TechTarget, Inc.

Accelerated Time to Value

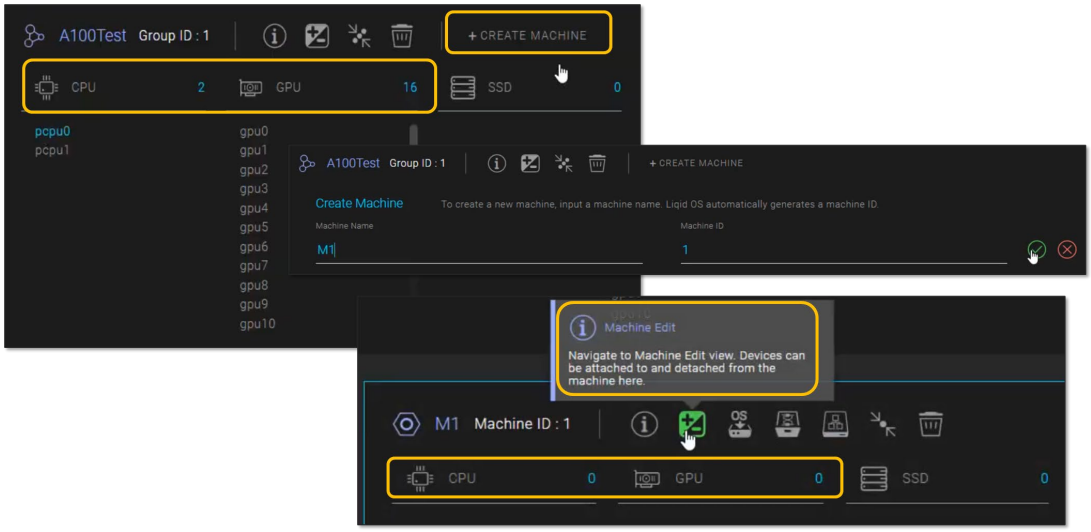
Configuring a server for a workload in minimal time is critical in today's business climate to meet business demands quickly and efficiently. However, the traditional process of provisioning and configuring server hardware (determining workload requirements; purchasing the required hardware components; then assembling, testing, and deploying the server) can take weeks or months to complete. With Liqid CDI, organizations can drastically simplify this process by configuring the servers on-demand via software, thus accelerating the time to meet business demands.

ESG Testing

ESG first configured a server requiring a large amount of GPU compute capacity to run a hypothetical artificial intelligence/machine learning (AI/ML) workload. We viewed the available hardware resources for configuring servers via the Command Center UI (see Figure 4), including two CPUs from our test bed, "pcpu0" and "pcpu1," and the 16 NVIDIA A100 GPUs, labeled "gpu0" through "gpu15." All components were assigned to a group[4] called ***A100Test.***

We began by clicking on ***Create Machine*** to compose an Ubuntu server and named it "M1." We noted that "M1" had no CPUs or GPUs assigned to it. We then clicked on the ***Machine Edit*** icon to assign resources to "M1."

Figure 4. Creating a New Machine "M1"



Source: ESG, a division of TechTarget, Inc.

ESG then assigned "pcp0" and all 16 GPUs to the "M1" machine by clicking ***Add*** next to each line item (see Figure 5). Recently added components were flagged by the green vertical line next to each component. After "pcp0" and the 16 GPUs appeared under the ***Assigned*** column, we clicked on ***Reprogram*** in the upper right-hand corner to compose "M1."

[4] A group enables an IT administrator to allocate specific resources to a user or team. The group can be restricted with Liqid's role-based access control (RBAC) functionality.

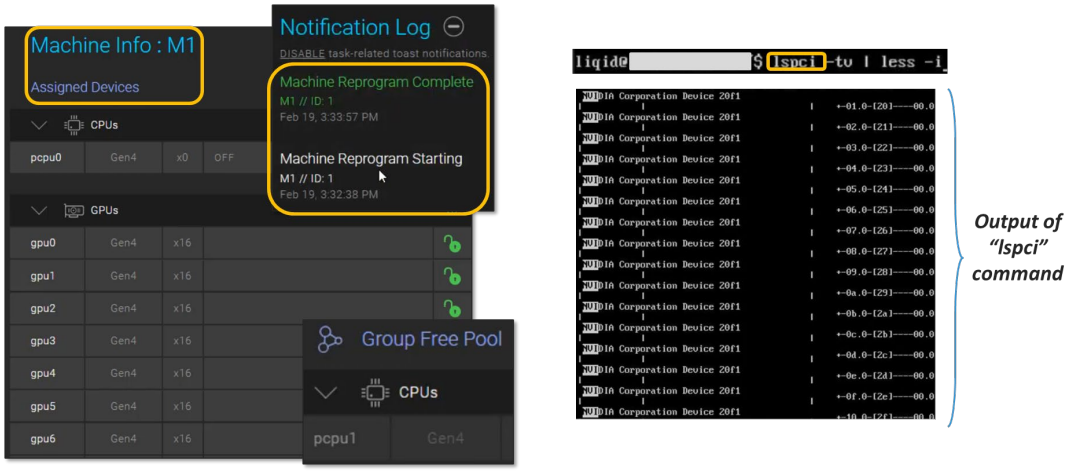
Figure 5. Adding Compute and GPU Resources to Machine M1

Source: ESG, a division of TechTarget, Inc.

ESG observed that the ***Reprogram*** step completed in under five minutes via notifications (see Figure 6). We also saw that these components were in the ***Assigned*** category when viewing the "M1" machine. In other words, "M1" was a bare-metal Ubuntu server connected electrically via PCIe Gen4 to the 16 GPUs, as if we physically installed 16 GPUs into a server. We also noted that the only resource available for configuration was "pcpu1," the second Intel Ubuntu server.

To confirm that the GPUs were connected via PCIe to "M1," ESG remotely accessed it and ran the "lspci"[5] utility. All 16 GPUs were listed as attached to "M1" and ready to be utilized.

Figure 6. Verifying that Available GPUs are Connected to "M1" Machine



Source: ESG, a division of TechTarget, Inc.

Observing the relative ease of configuring servers on-demand, ESG noted that the value of Liqid CDI goes beyond its ability to configure servers on-premises for any given workload in real time. By drastically shortening the IT provisioning cycle, organizations can immediately utilize data center resources and increase their return on investment (ROI) in less time.

[5] "lspci" is a command on Unix-like operating systems that lists detailed information about all PCI buses and devices in a given system.

Why This Matters

Configuring servers to support a workload's requirements is tedious and time consuming when using traditional IT provisioning processes. Ordering the proper hardware and additional components, configuring hardware, testing, and deploying new servers can last a few weeks to a few months. With this approach, organizations cannot meet their critical business needs in a timely manner. AS ESG research found, 67% of respondents agree or strongly agree that they are under pressure to accelerate IT infrastructure provisioning and deployment to support developers and line-of-business teams.

ESG validated that Liqid CDI accelerates time to value by enabling organizations to build and configure servers to meet precise resource requirements for any given workload in a drastically reduced amount of time compared to common IT provisioning cycles. We noted that a significantly shortened server provisioning process can help organizations to extract maximum business value from newly deployed servers in a minimal amount of time. Subsequently, deploying servers to run workloads almost immediately helps organizations to meet business demands quickly and efficiently, without having to schedule maintenance windows or refreshes.

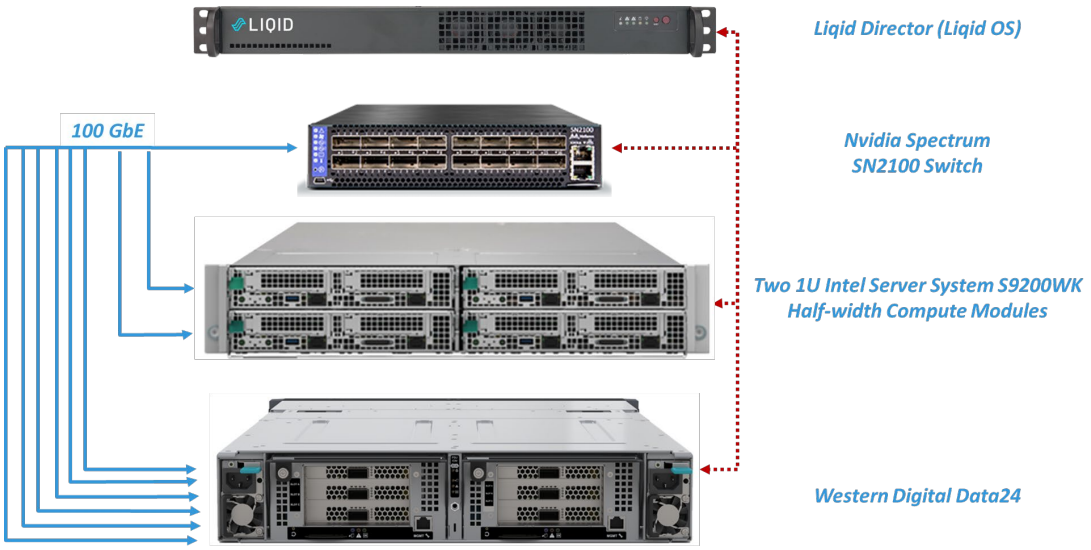
Increased Data Center Resource Efficiency

Organizations attempt to right-size servers for supporting and scaling new workloads. However, servers are overprovisioned at the onset to accommodate future growth, and resources are underutilized. When organizations scale or tear down workloads, a surplus of GPU compute and storage capacity can accrue. Reusing and redeploying this surplus is not an easy exercise. While organizations may attempt to "shoehorn" a workload onto a pre-configured server, that server either does not have sufficient compute or storage capacity to support the new workload's requirements, or the server has excess compute or storage capacity that will be wasted on a smaller, less demanding workload. With Liqid CDI, organizations can maximize utilization of data center resources in real time, freeing them up for other workloads.

ESG Testing

ESG used the testbed shown in Figure 7 to demonstrate Liqid's NVMe-oF composability. A 16-port NVIDIA Spectrum SN21005 100GbE switch was connected to two 1U Intel Server System S9200WK half-width compute modules via two 100GbE cables. Each server module contained an NVIDIA 100Ga ConnectX-6 SmartNIC. The NVIDIA Spectrum switch also connected to a single WD OpenFlex Data24 NVMe-oF storage platform via six 100GbE cables. The Data24 was populated with 24 Western Digital 3.6TB NVMe SSD drives. The Liqid Director was connected to the SN2100 switch to enable resources composability.

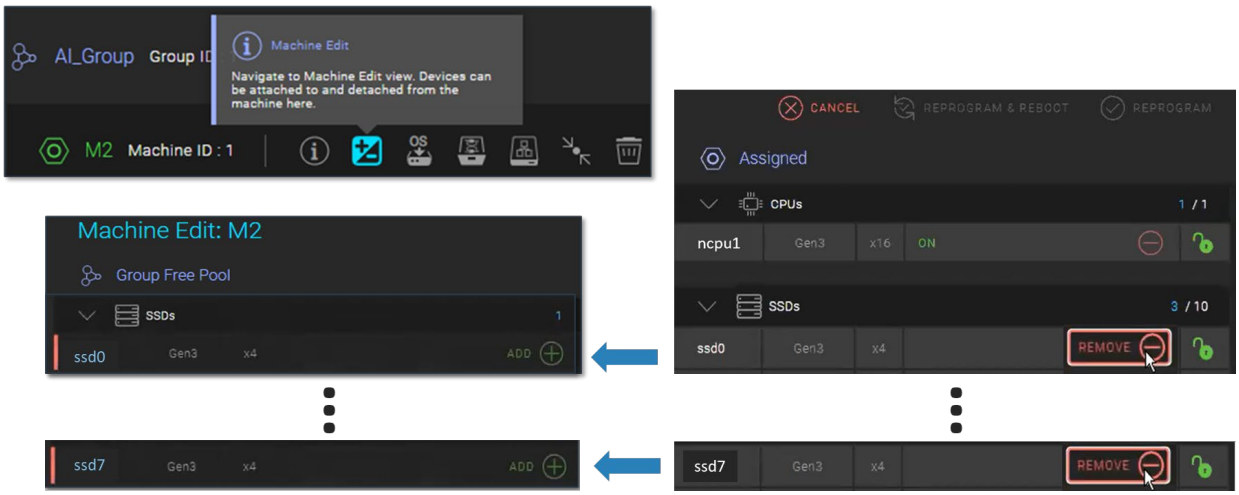
Figure 7. NVMe-oF Composability Testbed



Source: ESG, a division of TechTarget, Inc.

From the "M2" machine, an Intel Server System S9200K compute module containing the CPU named "ncpu1," ESG removed 8 Western Digital NVMe SSD drives located in the Data24 NVMe-oF enclosure, (see Figure 8). We first clicked on the ***Machine Edit*** button and saw these devices associated with "M2." We then scaled the server down eight NVMe SSDs by clicking on ***Remove*** next to "ssd0" through "ssd7." These devices were automatically reassigned to the ***Group Free Pool***.

Figure 8. Removing Excess Storage Capacity from "M2" Machine



Source: ESG, a division of TechTarget, Inc.

In this case, ESG saw how organizations can quickly free up resources from servers, such as when the workload no longer requires its existing storage capacity. Rather than allow the excess capacity to remain idle, Liqid CDI enabled us to move NVMe SSDs back to the free pool for use by other workloads.

ESG could also see how organizations can maximize GPU resource utilization by reclaiming underutilized GPU resources and assigning them to workloads needed for long durations. Organizations can also move them between daytime and nighttime workloads. For example, GPUs can be utilized for VDI during the day, then moved to a host that is performing ML training during off hours.

Why This Matters

Maximizing data center resource utilization is challenging when scaling or tearing down workloads. Existing compute and storage resources become available but reusing them is not straightforward. Either the excess GPU compute and storage capacity is wasted on a workload with less stringent requirements, or the capacity remains idle until moved to workloads that need it.

ESG validated that the Liqid CDI enables organizations to maximize data center resource utilization. By observing how simply we could remove Western Digital NVMe SSDs from an existing server configuration, we could see how storage resources are made available to be assigned to other workloads. Organizations can minimize capital expenses since these available resources can be reused by other workloads when needed.

Improved IT Agility

Modifying and scaling existing server infrastructure to accommodate changes in workload requirements incurs downtime, preventing any business from meeting demands in a timely manner. Organizations must navigate the IT provisioning process to either install additional components to an existing server or build additional servers to handle an increased load. Satisfying modified workload requirements will again require a few weeks to months, depending on the extent of the hardware changes to be made. On the other hand, Liqid CDI can help organizations to dynamically modify composed servers in real time, without the need to physically modify existing server infrastructure.

ESG Testing

Using the same test bed in Figure 7, ESG scaled up storage capacity using resources reclaimed from "M2" from the ***Group Free Pool*** to an existing machine named "M3." We first navigated to the ***Machine Edit*** page associated with "M3," an Intel Server System S9200K compute module containing "ncpu0" (see Figure 9).

Figure 9. Scaling Up Storage Capacity on Existing "M3" Machine

Source: ESG, a division of TechTarget, Inc.

ESG then added "ssd0" to "ssd7" by clicking on the ***Add*** signs for these eight line items under the ***Group Free Pool***. We then clicked ***Reprogram*** so that they would be connected via Ethernet to "M3." We also verified the addition of the eight Western Digital NVMe SSDS with "lspci."

ESG took note that the use of Liqid CDI not only eases the process of scaling storage capacity within an existing server but also increases the chances that the updated server configuration is implemented correctly. Manually installing additional

components into any server chassis is not difficult but does consume time unnecessarily and introduce unanticipated issues that prevent a successful server reconfiguration. With Liqid CDI, such errors are practically eliminated.

Why This Matters

To remain agile and responsive to business needs, organizations will need to scale certain workloads in real time, such as an AI/ML workload that must deliver real-time insights, even as the amount of data grows. However, scaling GPU compute and storage capacity in real time on-premises is simply not doable with today's IT operational processes.

ESG validated that the Liqid CDI helps organizations to modify and scale both GPU compute and storage capacity of existing on-premises servers in a matter of minutes with Liqid Command Center. We noted the simplicity of attaching additional Western Digital NVMe SSDs to a server in order to accommodate scaled-up workload requirements. We accomplished these tasks remotely without opening up an existing server chassis and manually installing components, helping to minimize operational expenses.

The Bigger Truth

Extracting the maximum value out of a server infrastructure in the data center is difficult when workload requirements change. While traditional IT provisioning processes enable organizations to meet workload requirements exactly, the risk of incurring unnecessary capital and operational expenses is high. If the workload scales up, more compute or storage capacity needs to be purchased and installed. If the workload scales down, excess capacity remains idle until another workload, with slightly mismatched GPU compute and storage requirements, can be supported. In both cases, capital and operational expenses increase over the long run, the ROI decreases, and the business cannot meet business demands in an efficient and agile way.

Liqid CDI enables organizations to compose servers on demand within on-premises data centers without wasting hardware resources. Leveraging pools of disaggregated GPU compute and storage devices interconnected over PCIe or Ethernet fabrics, organizations can compose servers via software to meet any workload's exact requirements. Once requirements change, Liqid CDI helps to modify server configurations with minimal delay. Should the workload be scaled back or torn down, components of the composed server are reclaimed by the resource pools, making them available for new workloads to reuse them. The average IT provisioning cycle can now last seconds or minutes, not weeks or months (assuming that the organizations have the proper number of components in the free pool to support new and existing workloads).

With Liqid's multi-fabric support, organizations can easily implement the NVMe-oF architectures that have been gaining attention for delivering high performance without incurring excessive costs. In fact, ESG research found that 66% are either interested in or planning to deploy NVMe-oF, while another 28% have already implemented this technology. The combination of Liqid's CDI and Western Digital's OpenFlex Data24 NVMe-oF Storage Platform enables organizations to quickly realize the benefits of composability with the storage performance of NVMe at lower costs.

Throughout our testing and evaluation, ESG validated that Liqid CDI can help organizations to compose servers on the fly that meet precise GPU needs, remove excess storage capacity from existing server configurations, and reassign those same components to another workload in a matter of minutes using Liqid Command Center. We observed that we could add any number of GPU and/or NVMe SSD devices to any existing server equipped with CPU and RAM, and scale them independently of each other. We verified that the devices were indeed attached to the server configurations we created using LINUX-based utilities.

ESG was impressed with the software-based solution Liqid offers to help organizations rethink their IT provisioning processes and uncover value from an on-premises data center infrastructure. We highly recommend taking a closer look at Liqid CDI should you want to achieve that end.

The goal of ESG Validation reports is to educate IT professionals about information technology solutions for companies of all types and sizes. ESG Validation reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objectives are to explore some of the more valuable features and functions of IT solutions, show how they can be used to solve real customer problems, and identify any areas needing improvement. The ESG Validation Team's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments.