Technical Insight Report

How Composable Infrastructure Addresses IT's Problem of Space and Time

By Eric Slack, Sr. Analyst

November, 2020. updated 2023

Evaluator Group

IT infrastructure plays a critical role in a company's success, and ultimately its survival, by enabling an enterprise to develop applications in the timeframe required to deliver what the market wants — before the competition does. To accomplish this, enterprise IT must solve a problem of space and time, where space is defined as infrastructure capacity and time as compute cycles. IT organizations have finite amounts of both resources and must maximize their use in order to assure they can deliver the compute environment the company needs when called upon.

Composable Disaggregated Infrastructure (CDI) is designed to address this optimization problem by enabling valuable resources to be deployed through software in just the right amounts and with a minimum of time between jobs or workloads. This report will examine how CDI address IT’s fundamental problem.

Space and Time

In an idealized scenario, the IT resources required to support an enterprise's applications would be deployed in exactly the right amount, at exactly the right time, and then redeployed to another application when no longer needed, with the same level of precision.

The total capacity (the space component) of the company’s IT estate would be fully utilized, and those resources would never be idle (the time component), due to set up, configuration changes, or expansion. While this scenario is never realized, its pursuit is more than just a goal. In a world where the high-performance resources that drive innovation are always in short supply, and the next breakthrough can never come too soon, solving the space and time problem is essential.

Integrated Systems

Integrated systems, hyperconverged (HCI), converged (CI) and CDI, have simplified the way IT builds infrastructure. They combine compute, storage and networking resources to create a comprehensive solution that provides agility, resource efficiency, and scalability, while improving productivity. The most common is HCI, a cluster of standard server and storage hardware that abstracts the storage resident in each node through a software-defined storage layer and virtualizes compute resources of each server node through a hypervisor, although HCI can also support a Kubernetes environment.

Software-defined storage (SDS) adds a layer of software to the stack, limiting the resources that can be abstracted and shared. The hypervisor adds yet another layer of software, impacting performance and increasing cost. And, by definition, the software-based architecture of HCI precludes the bare-metal connectivity that companies need for high-performance applications.

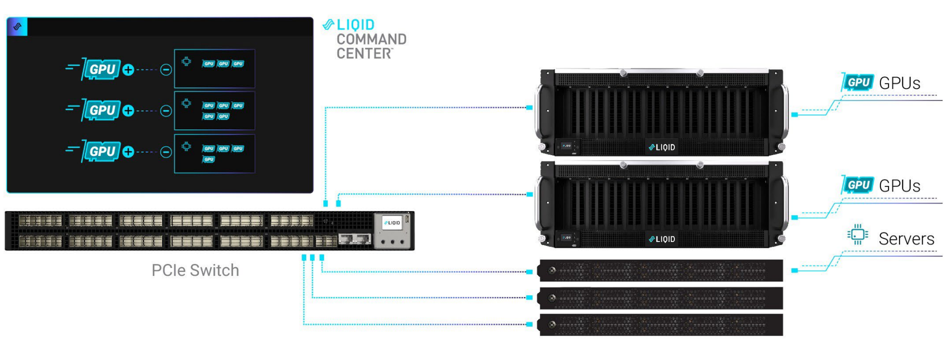
HCI doesn't share other resources, such as GPUs or FPGAs, that are needed in many high-performance workloads, and most don't have the disaggregated architecture required to separate those resources into pools for more efficient allocation. Finally, HCI relies on Ethernet to move data between nodes to create resiliency and to bring data and compute together.

Converged infrastructure is another approach to integrating storage and compute, adding networking to connect standard storage arrays and servers into a rack-level bundle. These systems provided bare-metal connectivity but don't allow for the simple reallocation of those resources to support a dynamic set of workloads.

CDI combines some of the best parts of Converged and Hyperconverged technologies, extending the benefits of abstraction to the hardware. Instead of inserting a software virtualization layer like HCI does, composable maps disaggregated, bare-metal resources to compute modules (CPU and RAM) to create a high-performance server instance that can be reconfigured as needed, without any physical changes.

Composable Disaggregated Infrastructure

Flexibility is a primary motivation behind Liqid's decision to create a CDI solution that's based on a PCIe fabric (see figure below). This fabric is built around a managed PCIe switch, with up to 48 ports of PCIe Gen 4.0 x4, x8, or x16 lanes. This architecture allows any PCIe-compatible device to be connected into the fabric and composed. Most often these are standard expansion chassis for storage - SAS or NVMe SSDs or Storage Class Memory (SCM) devices - or PCIe Add-in Cards that contain accelerators, like GPUs or FPGAs, each of which can be composed at the device level.



PCIe Fabric CDI (Image courtesy Liqid, Inc.)

Liqid offers PCIe expansion chassis that support up to eight Full Height, Full Length PCIe Gen4 add-in-cards providing 100ns of fabric latency and 64 GB/s port bandwidth in a hot-swap design. Of course, the Liqid PCIe fabric supports any compatible devices in any PCIe expansion chassis. The result of this open architecture is that more total resources can be composed, with the ability to scale the fabric to support thousands of devices.

Composing Software

The Liqid Command Center management software composes granular pools of resources into bare-metal systems within a few seconds. These systems are custom configured for the workloads they will support. Command Center, which runs on a low-cost server or embedded into a switch connected to the fabric, orchestrates this process at the device level through a GUI, CLI, or a set of RESTful APIs. In addition, this software manages scaling of resource pools, multitenancy, remote configuration, migration, etc.

Scalability and Resource Options

The Liqid systems do not have capacity limits imposed by the number of slots in a server chassis. There is also no restriction to specific configurations of nodes that the vendor offers in terms of CPU speeds, core counts, number of sockets, SSDs supported, etc. In fact, there is no limitation to use resources from any vendor. This architecture also enables the use of storage, accelerators, and networking cards from any manufacturer. There’s no lock-in to Liqid or any vendor that sells PCIe cards, SSDs, GPUs, etc. Furthermore, Liqid’s fabric can be extended to include existing rack servers or blade servers.

Liqid’s fabric architecture provides resource granularity as well. Any PCIe SSD, from any vendor can be used, giving the full range of capacities available on the market. Liqid offers its own storage options that include PCIe Add-in-Cards comprised of multiple NVMe or Storage Class Memory modules, with each module individually composable. The bottom line is, when compared with a blade server architecture, Liqid PCIe fabric-based systems allow more flexibility for the composing process, more total capacity, more types of resources and a choice of vendors.

Flexible Refresh Cycles

Disaggregation is the term that describes the separation of physical resources, most often compute, storage and GPU, allowing them to be efficiently pooled and allocated. This could be called a "technical" benefit, where better resource efficiency is a result of disaggregation. There is, however, another benefit, one of organizational efficiency and productivity that’s also driven by resource disaggregation.

Servers are typically refreshed every 3-5 years, based on the CPU refresh cycles that produce more compute capacity and performance. The ability to separate storage from compute, also separates the buying cycles of servers and storage devices from their respective lifecycles. This enables companies to upgrade each resource at the optimal time, more frequently when performance is critical and less frequently when cost is more important.

Storage devices within a PCIe expansion chassis are hot-pluggable and can be upgraded at any time, independent from other storage resources or the CPUs. There is no requirement to wait for the chassis vendor to release a new storage blade to take advantage of an upgrade in flash technology or a larger capacity SSD.

Summary

Composable is a technology that relies on flexibility. Composable systems should provide the most options with which to use that flexibility; meaning more resources, like FPGAs, GPUs, NICs and storage-class memory. A system that only composes storage is counter to the composable concept.

Disaggregation is another key component of composable technology, providing the means to pool more resources for increased efficiently and scale. Central to leveraging this functionality is the ability to compose resources in different form factors, from different manufacturers.

Evaluator Group Comments

Converged and then Hyperconverged Infrastructures were the first integrated systems, created to simplify the design and deployment of IT infrastructure and improve scalability. But they didn't have the granularity or device- / hardware-level control to efficiently allocate resources, all the resources, needed to address IT's problem of space and time.

Servers have a finite number of PCIe slots. For applications that need GPUs, like advanced analytics or high-performance computing, this creates an upper bound on capacity, a limit on the number of applications that an infrastructure can support and a limit on the time which they can be run.

This makes the ability to compose those GPUs more than just interesting, *it makes that capability compelling*. With a PCIe fabric architecture, Liqid takes the capacity limit off compute, storage and high-performance accelerators, and extends the compute time available for critical applications, helping IT solve the problem of space and time.

This report was produced with funding from Liqid, Inc.

About Evaluator Group

Evaluator Group LLC. A division of The Futurum Group, ia an Information management and data storage analyst firm, has been covering systems for over 20 years. Executives and IT Managers rely upon us to help make informed decisions to architect and purchase systems supporting their data management objectives. We surpass the current technology landscape by defining requirements and providing an in-depth knowledge of the products as well as the intricacies that dictate long-term successful strategies.