# Composable Infrastructure – Datacenter for the Next Decade

## Executive Summary

*In this Insight, Neuralytix examines and analyzes the next generation of datacenter infrastructure architecture – Composable Infrastructure. Neuralytix believes this will be the next major datacenter infrastructure architecture that will be adopted by larger enterprises in the next decade.*

*The architecture may seem revolutionary, but it is essentially the disaggregation of converged and hyperconverged infrastructures (CI and HCI).*

*Neuralytix defines composable infrastructure, or Datacenter 4.0 (DC4), is a datacenter infrastructure that can be programmatically software-defined using disaggregated industry-standard components to support a wide variety of application workloads.*

*Vendors including Dell EMC have already begun shipping DC4, while others such as HPE are shipping a proprietary version of DC4.*

## Contents

## Introduction

In this Insight, Neuralytix will introduce the concept of Composable Infrastructure, which we also refer to as Datacenter 4.0 (DC4), and why this type of infrastructure is the future of infrastructure architecture for 2020 and beyond.

*Composable Infrastructure, or Datacenter 4.0 (DC4), is a datacenter infrastructure that can be programmatically software-defined using disaggregated industry-standard components to support a wide variety of application workloads.*

## In the beginning …

In the beginning, there were minis and mainframe computers. These were scalable computers primarily had a single or a small number of applications running on them.

Both minis and mainframes were scalable (albeit, in a limited fashion). Mainframes, in particular, allowed enterprises to scale up the number of computer processors, memory, and storage. Since mainframes were self-contained, networking was not really an issue.

Minis changed that by allowing applications to run across a number of mini computers (such as the DEC VAX), and introduced the first storage area network (SAN), such as the DEC CI cluster.

Both of these types of computers can be considered centralized, since they were neither portable, and required special flooring, and dedicated computer rooms to function optimally.

With the arrival of the x86 based PC servers, rather than centralizing functions, special or specific tasks (or applications) could be distributed amongst many servers interconnected by a network interface.

PC servers grew in popularity, and has since spawn servers all around the world interconnected not only by private networks, but by the Internet as well.

Despite this proliferation, one thing that could not be changed was the need to aggregate data and information. This meant that data had to be brought back into a centralized data center for the data and information to be processed.

This led to the beginning of monolithic storage systems such as those from EMC, HDS, and IBM. These monolithic storage systems, could hold large amounts of data, but required the customer to commit to the technology for a long time since, they would be investing in a system that would be large physically, but small in capacity, and they would fill up that capacity over time.

These systems were not suitable for smaller businesses, and the modular storage system was born.

By the current decade, everything had become modular, but the same proliferation of PC servers that was noted earlier has now become an outbreak of PC servers, augmented by hypervisors, created a seemingly endless number of virtual servers.

By the mid part of the 2010s, a new type of infrastructure had emerged, hyperconverged infrastructure (HCI). The most basic principle for HCI is simplicity in management and scale. Essentially, enterprises would be buying units of IT that aligned with their necessary workloads. When they needed more, they could scale the number of HCI nodes in parallel with business needs.

This brings us to the modern HCI architecture, and the emergence of DC4, and why the future is in composable infrastructures and architectures.

To be fair, many HCI vendors already have products that can be easily adapted for DC4. So, this change will be witnessed more as an evolution than a revolution.

# An examination of the infrastructure architectures of the past

Also, our narrative above gives a brief qualitative description of the evolution of the datacenter infrastructure story, below, we give a more graphic and more detailed explanation and description of each generation.

We start from the post minis and mainframe era, but pre-PC server timeframe.

## Datacenter 1.0 – Traditional Infrastructure

*Figure 1: Datacenter 1.0 - "Traditional Infrastructure" (Neuralytix 2017)*

Traditional infrastructure is often called client/server architecture, we saw high end specialized servers with unique processors, such as the Sun Enterprise 10000 in Figure 1. These systems were married with monolithic storage systems such as EMC Symmetrix, and connected over an Ethernet work using director class networks such as those from Cisco.

Traditional infrastructure started from around 1995, and many enterprises still run traditional infrastructures architecture due to legacy applications.

While PCs would often sit on the same Ethernet network, data essentially came from clients, which were then uploaded or inputted into these big servers, that ran enterprise databases and applications such as the Oracle database, and SAP ERP systems.

The specialized servers would process the data, and the result would return to the clients. In a small evolution of traditional client/server, there was also three-tier client/server architecture that not only had clients and (typically) server, but there was a database server that was distinct from the application server.

Nonetheless, the architecture is the same.

### Datacenter 1.1

At the same time as the traditional infrastructure leveraging highly specialized CPUs (such as Sun SPARC and IBM POWER), PC servers also became popular, particularly for smaller applications or workloads and for smaller enterprises. PC servers could also connect to monolithic storage, and director class Ethernet switches.

But comparatively, PC servers were not expected to match or exceed enterprise servers in performance or scale. Yet, Intel continued to develop more and more powerful CPUs, and enabled multi-socketed, multi-core configuration, that by 2010, x86 based servers were as powerful, if not more powerful, than enterprise servers.

## Datacenter 2.0 – Reference Architectures (RA)


*Figure 2: Datacenter 2.0 - Reference Architectures (NetApp 2017)*

One of the largest challenges facing customers with respects to traditional infrastructure (aka DC1.0) was interoperability.

In 2010, NetApp, in conjunction with Cisco and VMware, launched FlexPod, a reference architecture

(RA) to rival VCE's VBlocks (see next section on Datacenter 2.5).

These RAs were essentially "cookbooks" or "recipes" by which certain configurations were pre-tested and found to meet some certain criteria of compatibility. In essence, these solutions were certified bundles that could be bought from a reseller with a single SKU. This resulted in a more predictable level of compatibility, and dramatically improved the simplicity of purchasing.

However, the challenge for the customer is that it needed to essentially purchase a FlexPod "solution" from three different vendors – Cisco, NetApp, and VMware, and carry three separate support contracts.

So the purchase was easier, as the reseller would abstract the deal down to a single SKU (or a single deployed solution), but the challenge of support and licensing continued to be a challenge.

## Datacenter 2.5 – Converged Infrastructure (CI)


*Figure 3: Datacenter 2.5 - Converged Infrastructure (EMC 2016)*

As PC servers gained momentum, spurred on not only by the performance and capabilities (such as being able to address larger and larger memory pools, with multiple layers of cache), the equivalent PC server compared to an enterprise server was at least an order of magnitude less in cost.

In 2009, VCE (which started its life as Arcadia), a joint venture between Cisco, EMC, Intel, and VMware delivered a datacenter that was fully tested, warranted, and maintained. This is known as converged infrastructure (CI).

Companies like VCE would bundle compute, network, storage, and management as a single SKU, for easy purchasing, and had a technical support team that required a single support contract.

CI had the architectural benefits of RA, such as guaranteed compatibility, and simplicity of purchase, but it went one step further by providing integrated management and a single support contract for all the components.

Large corporations were enamored by this as it created a standard which produced a predictable outcome, lowered risk, while getting best of breed components.

## Datacenter 3.0 – Hyperconverged Infrastructure (HCI)


*Figure 4: Datacenter 3.0 - Hyperconverged Converged Infrastructure (Nutanix 2017)*

For smaller corporations, while the attraction of VBlocks and CI, the entry points were not as attractive.

Start-up companies such as Pivot3, Scale Computing, and eventually Nutanix delivered the same architectural promise as CI, in a smaller form factor, and at a much lower price point.

We call this Datacenter 3.0 (DC3.0) or more commonly, hyper-converged infrastructure (HCI). HCI essentially brought all the components of compute, memory, storage, and networking, and abstracted them through a hypervisor, software-driven or software-defined storage (SDS), and software-defined networking (SDN) into a single software-driven or software-defined datacenter (SDDC). SDDC is synonymous with HCI, and software-defined infrastructure (SDI).

These nodes of compute, memory, storage and networking also had the benefit of scale-out capabilities allowing companies to start small, and scale as their needs required.

HCI is commonly deployed for virtual desktop infrastructure (VDI), as well as tier 1 business and mission critical applications, including, but not limited to Oracle, Microsoft SQL, etc.

HCI allowed many corporations to evolve from their traditional infrastructure (DC1.0) to a modern flexible, scalable, predictable architecture, without a rip and replace strategy. Still some more enterprises that had taken the first steps towards convergence through RAs and CIs, are also moving towards HCI.

One main critique of early HCI is that it does not properly scale compute and storage capacities independently. While this can be true of some HCI solutions, this has already been addressed by many HCI vendors.

## Datacenter 4.0 (DC4) – Composable Infrastructure


*Figure 5: Datacenter 4.0 (DC4) - Composable Infrastructure (Liqid, Neuralytix 2017)*

While the modular and scalable aspects of HCI are attractive to many organizations, very large enterprises, as wells as service providers want to see even more flexibility.

This is where composable infrastructure, or what Neuralytix terms DC4 comes into play. This is the next generational datacenter in which compute, GPU, memory, storage, and networking are all commoditized and disaggregated, and through *software* only, are the components brought together (or *composed*) to form composed physical servers (bare metal machines).

Every resource of compute, GPU, memory, storage, and networking are essentially considered independent components, and this gives rise to a very flexible, and highly dynamic operating environment that can be deployed and decommissioned via code. As such, DC4 is sometimes referred to as Infrastructure as code.

There are current technical limitations in DC4 that are constraint by the laws of physics. However, Neuralytix forecasts that by 2020, roughly 5% of datacenters will be running some form of DC4 architecture, making DC4 a material datacenter architecture of the future.

That said, Neuralytix also believes that HCI and DC4 will coexist for at least a decade as they serve differentiated outcomes, and satisfies differing views on datacenter architectures.

## What is Composable Infrastructure (DC4)?

Composable Infrastructure, or Datacenter 4.0 (DC4), is a datacenter infrastructure that can be programmatically software-defined using disaggregated industry-standard components to support a wide variety of application workloads.

The DC4 definition is a forward looking one. The definition of DC4 calls for the aggregation of disaggregated industry-standard components, which includes CPU and memory.

That said, it is important to remember that DC4 is designed to support a wide variety of workloads. Therefore, workloads that require extreme amounts of CPU and memory should turn to other datacenter infrastructures, such as HCI or CI, to fulfill those needs. Alternatively, applications will need to be redeveloped to operate in a clustered environment.

Figure 6 shows a logical representation of how an administrator can *compose* physical machines from available pools of compute, GPU, memory, storage and networking resources, simply by dialing up or down the amount of resources needed per machine.

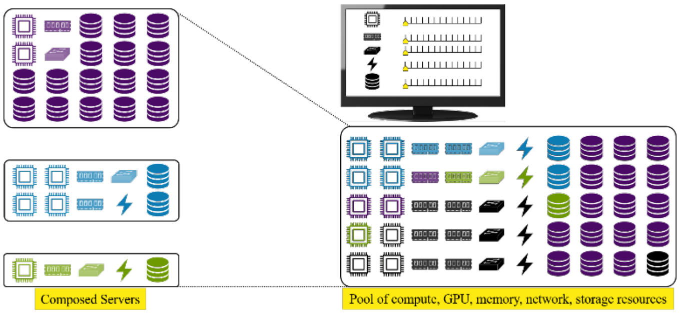


*Figure 6: Composing Physical Servers (Neuralytix 2017)*

On the left-hand side of Figure 6, are three machines that have been composed. The top machine is heavy on storage, and could represent an archiving server. In the middle is a compute heavy machine, which would be suitable for running a database. Finally, the machine on the bottom is a general purpose server, that has a relatively small amount of compute, memory, network and storage resource, and as such could represent a virtual desktop infrastructure (VDI) session. Industry standard hypervisors can be deployed on top of composable infrastructure environment to enable more power and flexible virtualization layers.

For each of these scenarios, an administrator would simply adjust the sliders to dial-up or dial-down the required resources for a given business need.

Furthermore, these resources, per the definition of DC4, can be programmatically adjusted through code (sometimes referred to as Infrastructure as Code). This allows for a great deal of automation, and compliance to policies either driven by business or regulatory needs.

# What you can and cannot do with DC4

Earlier in this section, we noted that the DC4 definition is a "forward looking" definition. What we mean by that is that it considers capabilities that are not yet available. In the near term, we expect that DC4 will be deployed with CPU and memory as a single assignable unit until next generation technolgoies enable the true dissagregation of memeory.

## What works

Composable infrastructure is all about the programmable aggregation of disaggregated components.

Software-driven (or software-defined) networking and storage is possible today, and has been available for a long time. Both technologies are mature and well understood.

Disaggregation of compute resources are possible via software such as hypervisors (kernel mode virtualization) and containers. Again, both are well understood, although containers are still maturing.

In both cases, software is used to aggregate and disaggregate hardware infrastructure as required.

## What does not work

What is not working well yet is the aggregation of memory resources and CPUs. Aggregation of memory resources *may* be possible in the near future, but only when memory class fabrics standards such as Gen-Z (currently being proposed) exist and become commercially viable.

Equally, aggregation of CPUs and cores across motherboards *may* be possible in the future, also via software from some emerging vendors. Aggregation of CPUs and cores challenges the laws of physics. So long as the communications between the cores and layer 1, 2 and 3 cache require proximity to achieve minimal or no latency. Once signals leave the physical motherboard, additional protocols and distance introduces latency and new complications which may reduce the viability of such technologies. Startups such as TidalScale are already trying to make such technologies commercially viable.

A question as it relates to cross-motherboard begs to be asked, why do this at all? The argument would suggest that we could use composable infrastructure to create a high-performance computing (HPC) environment. But we have already solved this problem by having software that breaks down jobs and sends them in parallel to different processors and servers. (The Hadoop framework is one such way).

Additionally, if an organization desires to use a HPC environment, why take a general purpose infrastructure architecture to address a highly specialized problem?

This said, Neuralytix believes that the CPU/core aggregation problem will be solved within the next 5 years.

## Extensibility

One of the more interesting aspect of DC4 is its ability to be extensible. Apart from the ability to push software extensions and capabilities to individually composed server/desktop, it is also possible to incorporate hardware extensions and capabilities (see Figure 7).

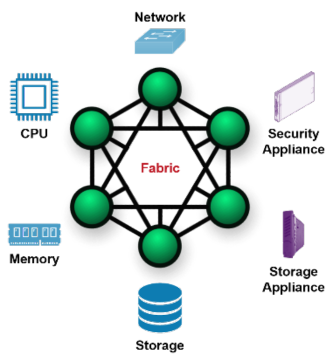


*Figure 7: Extensibility of Composable Infrastructure (Neuralytix 2017)*

Using low latency and RDMA over Fabric technologies such as RDMA over Converged Ethernet (RoCE) or iWarp, it is possible to bring specific capabilities that are hardware accelerated into a DC4 architecture.

An example of such is illustrated in Figure 7, where a/some security appliance/s and/or a/some storage appliance/s can deliver hardware accelerated security capabilities (e.g. hardware accelerated virus protection) or storage capabilities (e.g. hardware accelerated data reduction) to some or all of the data running on DC4.

These capabilities can be deployed over composed components. This is in contrast to CI and HCI, where security and/or storage appliances would have to be deployed against specific physical disk drives or servers.

This extensibility is also scalable. As the needs for hardware accelerated appliance increases (e.g. if there is a high amount of data reduction that needs to be done), more and more appliances can be introduced as additional resources into the composed servers that require it.

The extensibility of a DC4 architecture is theoretically limitless. Hardware appliances designed to do protocol translation, object to file/block conversions, or even format conversions could be introduced.

This is part of the flexibility and granular control that DC4 architectures bring to the table. Multiple similar composed environments could be created simultaneously, and managed collectively; or, individually composed, very specific could be created, with highly specialized hardware and software extensions.

## Guidance and Conclusion

Neuralytix advises that DC4 is not for everyone. We believe that DC4 is best suited to very large companies or service providers, where servers need to be brought up and decommissioned on a regular basis through automation. For those applications that require a dedicated set of resources, DC4 is not likely to be the best approach, and should be deployed through the use of CI and HCI.

Another point to note is that Neuralytix rejects any DC4 approaches that uses *any* proprietary or vendor specific components. This may include, but is not limited, vendor specific backplanes, custom protocols, etc. All resources should be addressable, and programmable through open and industry-standard APIs.

## Market Size and Forecast

For the calendar years 2017 and 2018, Neuralytix believes that there will be a *de minimis* amount of demand for DC4 (See Figure 8).

However, as we enter 2019-2021, Neuralytix currently forecasts that by 2021, the market for DC4 will approach $2 billion, and represent just under 5% of the datacenter spend worldwide.

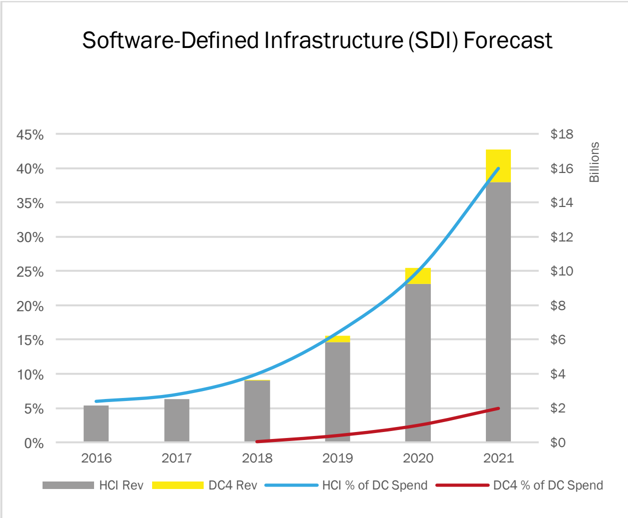


*Figure 8: Software-Defined Infrastructure (SDI) Forecast (Neuralytix 2017)*

This in contrast to the over $15 billion market and 40% penetration of HCI. While outside the formal forecast period for Neuralytix, we believe that the demand for DC4 is not likely to exceed 20% of the datacenter market through 2030.

Neuralytix also believes that customers of DC4 are likely to buy complete end-to-end systems from a single vendor, despite the open and industry-standard nature of DC4. This is because DC4 is about simplicity and flexibility. While the attraction of a do-it-yourself (DIY) approach to both building a DC4 architecture may be appealing, the maintenance and support of the infrastructure is key. Just like HCI updates are all-encompassing – covering compute, network, and storage – DC4 upgrades will affect compute, GPU, memory, network, and storage also, and a lack of synchronicity, and compatibility to result in less than ideal reliability.

The difference between HCI (or datacenter 3.0) and DC4 is the "as-a-Service" part of Infrastructure-as-a-Service (IaaS). A big part of DC4 will be the self-service capabilities of the software that drives it. Once implemented, the customer experience should be as simple (or simpler) than Amazon Web Services (AWS), Google Cloud Platform (GCP), and Microsoft Azure; but rather than it being a cloud-based infrastructure, DC4 will be on-premise, or collocated.

Service providers may even consider building compatible stacks that would physically sit next to cloud stacks so that customers can truly have a hybrid experience without loss of control.

Notes