Hewlett Packard Enterprise

Reference Architecture

HPE Reference Configuration for client virtualization in healthcare on HPE Synergy Composable Infrastructure

Delivering optimal user experiences on demand with Citrix Cloud

Contents

Executive summary 3
Solution overview 3
HPE Synergy Composable Infrastructure 5
HPE 3PAR 8440 unified All-Flash Array 5
NVIDIA Tesla T4 GPU 5
Citrix Cloud 5
Citrix Content Collaboration 5
Citrix Virtual Apps and Desktops Service 5
Solution components 10
HPE Synergy 10
HPE Synergy 12000 Frame 13
HPE Synergy Composer 2 13
HPE Synergy Image Streamer 14
HPE Synergy 480 Gen 10 Compute Module 14
PCIe graphics expansion module 15
Benefits of HPE Synergy Composable Infrastructure for end-user computing environments 15
HPE 3PAR 8440 AF 17
NVIDIA GRID 19
NVIDIA Licensing 20
NVIDIA Tesla T4 GPU 21
Software Components 22
HPE OneView 22
Citrix Cloud 22
Citrix Virtual Apps and Desktop Service 23
Citrix HDX 3D Pro for GPU based medical imaging applications 23
Citrix Content Collaboration 23
Summary 25
Appendix A: HPE Synergy composable infrastructure benefits in VDI/virtual environments 26
Appendix B: Bill of materials 27
Resources and additional links 29

Executive summary

As the Healthcare Industry grows and becomes more digitally connected, the IT teams that support healthcare organizations must find better ways to leverage technology to improve patient outcomes, work experiences for physicians and staff, and business operations. Integrating new, future-ready technologies will help to ensure the right experience for each end user (patients, doctors, nursing staff, etc.). The challenge is to deliver this experience in a cost-effective way with the right mix of technology.

End user computing has maintained a strong foothold in the healthcare industry for well over a decade because it addresses many of the challenges listed above. However, there is still a need to understand where the power of graphics processing units (GPUs) can add the most cost effective value in the technology experience for every user. While adopting GPU everywhere could indeed provide a quality experience, not every user needs a GPU. Moreover, deploying GPUs for users that do not need them is expensive. It is crucial to address the needs of a variety of workers (with and without GPU requirements) in a healthcare setting.

This Reference Configuration guide will show how to use HPE Synergy, Citrix® Virtual Apps, and NVIDIA® to achieve the outlined goals. HPE Synergy offers the flexibility to provision and deliver the right mix of resources – both non-accelerated (non-GPUs) and accelerated virtual apps with GPU. This solution gives workstations "as-needed" access to graphical applications, but does not require virtualized GPUs (vGPUs) in workstations and graphical acceleration.

Working with Citrix and NVIDIA, we can present:

- Virtual desktops and applications
- Accelerated applications
- Accelerated desktops

HPE Synergy can uniquely deliver this solution with the highest user density for virtual desktop infrastructure (VDI) and server delivered applications. To reiterate, more VDI sessions per rack for medical applications and electronic health records (EHR) than any other solution without sacrificing security, provides the best user experience possible while maintaining your data security.

Other unique security innovations on HPE Synergy include the Silicon Root of Trust technology that extends across the lifecycle for end-to-end security. HPE 3PAR enables hosting user data, hosting VMs that form the solution stack and allows scaling.

This Reference Configuration defines the best cost-effective solution healthcare organizations can deploy to deliver the right experience for their end-users. This document demonstrates the following:

- The benefits of running accelerated GPU workloads in EUC environment, leveraging the strengths of HPE Composable Infrastructure, Citrix, and HPE 3PAR.
- The introduction of "as-needed" consumption for high-performance GPUs to accelerate the applications
- GPUs repurpose display to compute mode to address both visualization and analytics workloads
- Composable architecture flexibility for healthcare related workloads on a single platform

Target audience: This document provides recommendations for configuring EUC solutions for healthcare with HPE Synergy platform. This will educate IT decision makers, enterprise architects, and partners and provide understanding about the groundbreaking HPE Composable Infrastructure, which enables you to compose fluid pools of physical and virtual compute, storage, and fabric resources into any configuration for diverse workloads.

Solution overview

This Reference Configuration highlights the value of HPE Synergy Composable Infrastructure, GPUs, and the Virtual Apps and Desktop service in Citrix Cloud and defines how this combination delivers a well-designed client virtualization solution for healthcare organizations.

High-end graphics applications and GPU enhanced artificial intelligence and inferencing applications continue to change the nature of healthcare. In addition, video and other graphics technologies need graphics-acceleration, making GPUs a core component for a healthcare solution. On the other hand, practitioners spend most of their time on text based electronic health record (EHR) which does not need high-

performance graphics acceleration. Hence this solution which is designed to optimally address computing needs of both text and graphics workload.

Figure 1 describes the healthcare system use cases with percentages

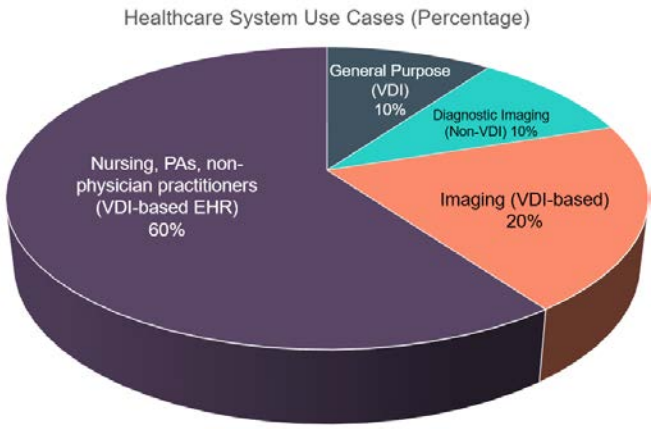

Figure 1. Healthcare system use cases

Architecting a right sized VDI solution for healthcare includes understanding the usage pattern and workload types. As shown in Figure 1, this has only a limited number of users who would use accelerated applications and hence this solution is designed to virtualize the high-end graphics applications instead of running them on individual user VMs. In this way, high-performance GPUs can be deployed where the application will use them most effectively and the GPU resources are used on-demand instead of being pre-allocated to individual desktops. To summarize, application virtualization provides access to high-end GPUs at far less cost due to the efficiency of an on-demand model and these lower costs are translated to lower healthcare cost per capita.

Figure 2 illustrates the healthcare system architecture diagram and provides the outline for delivering non-graphical and graphical applications across a healthcare organization, which includes EHR traffic, accelerated 3D graphics traffic and DICOM/H7 traffic.

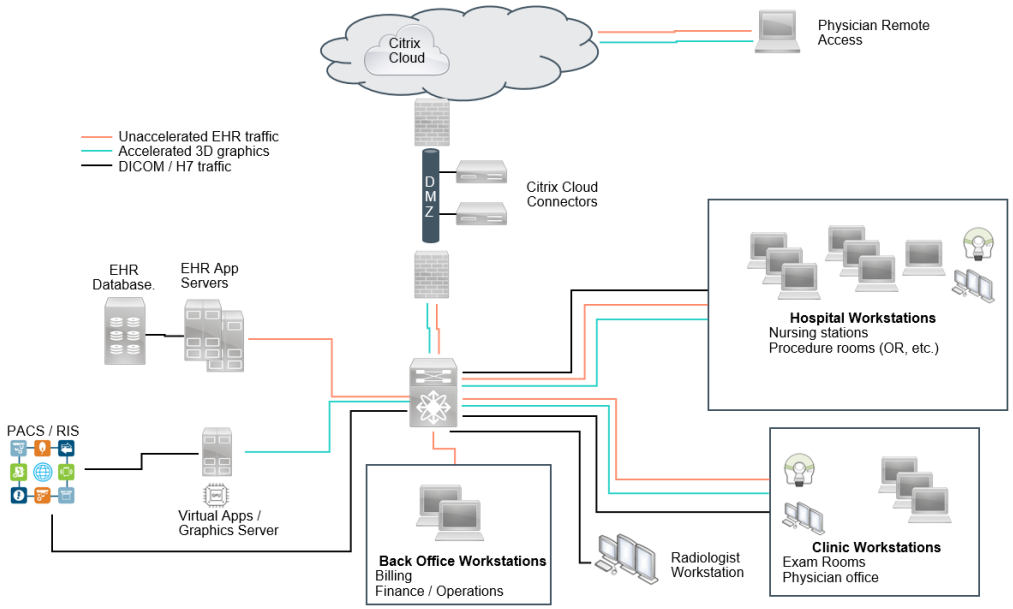

Figure 2. Healthcare system architecture

This Reference Configuration provides the solution for client virtualization to deliver quality and enhances end users experience, high performance and increased user density while improving security and manageability.

HPE Synergy Composable Infrastructure

HPE Synergy Composer appliance powered by HPE OneView manages the HPE Composable Infrastructure. HPE Synergy can compose and decompose compute, storage, and network to the respective servers on an on-demand basis. HPE Synergy Image Streamer provides stateless VMware ESXi hypervisor compute nodes for the VDI environment, so that the downtime can be minimized in case of any hardware failure. The solution comprises of HPE Synergy 480 Gen10 Compute Module with PCIe expansion module hosting NVIDIA GPUs to facilitate the Citrix Virtual Apps and desktop infrastructure.

HPE 3PAR 8440 unified All-Flash Array

The solution is built around HPE 3PAR unified All-Flash Storage that provides a unified storage solution to host the virtual apps and desktops and file shares. HPE 3PAR All-Flash Array provides accelerated read and write performance and low latency thus providing great application response time at peak usage and boot storms. The solution is designed aptly to host the healthcare consumers medical data, diagnostic reports, and high quality medical images and EHR database on HPE 3PAR 8440 All-Flash unified Storage Array as file data is managed and administered using HPE 3PAR File Persona. Since the file data resides on All-Flash disks, the read and write of file data provides excellent response time to end users.

NVIDIA Tesla T4 GPU

Physicians depend heavily on accessing the diagnostic medical scan reports of patients and healthcare consumers. The solution has been designed precisely to address the challenges of user experience by healthcare personnel. HPE Synergy with NVIDIA Tesla T4 GPU provisions vGPU to Citrix Virtual Apps and Desktops, these accelerated applications and desktops are accessed by medical staff who benefit from precise and detailed view of the medical images backed by the NVIDIA GPU and Citrix HDX 3D Pro technology thereby providing doctors & medical staff a high quality image.

Citrix Cloud

Citrix Cloud is a cloud-based control plane of service offering and it contains management components for the Citrix Virtual Apps and Desktops service as well as Content Collaboration. The design shows a demarcation of Citrix components residing in the data center and Citrix Cloud. The Citrix Cloud Connector is the component that enables the connectivity between resource locations hosting the Citrix Virtual Desktop Infrastructure on HPE Synergy and the Citrix Cloud control plane. It is installed on two or more servers per resource location for redundancy and performance.

Citrix Content Collaboration

Citrix Content Collaboration enables end users to access the collaborative medical data and content securely from anywhere. The solution allows doctors and healthcare staff to access content within the organization. At the same time, this solution also enables outpatient users, mobile healthcare personnel and mobile doctors by letting them access data from anywhere thus providing a holistic approach for end users to access content from anywhere and from any device.

Citrix Virtual Apps and Desktops Service

End Users can have Virtual apps and desktops provisioned on HPE Synergy and access their respective desktops, medical data and applications via endpoints. Citrix Machine Creation Services (MCS) is used to provision accelerated or non-accelerated virtual applications for end users to manage and maintain EHR of healthcare consumers, view Digital Imaging and Communications in Medicine (DICOM) images and it provisions Virtual Desktops with or without vGPUs for doctors, medical staff and administration staff for performing their day-to-day activities.

Figure 3 highlights the solution components used to achieve the above-mentioned benefits and provides an architectural overview of the client virtualization solution. This solution is designed with nine HPE Synergy 480 Gen10 Compute Modules. The first four modules feature four NVIDIA T4 GPUs each: three provide enhanced visualization for published applications and one is dedicated for accelerating virtual desktops. The remaining five modules do not feature GPUs, which caters to the majority of healthcare workers including those that interact with EHR systems. Virtual machines, EHR data and other healthcare centric information are hosted on HPE 3PAR 8440 StoreServ Storage. Citrix Content Collaboration enables collaboration between healthcare personnel.

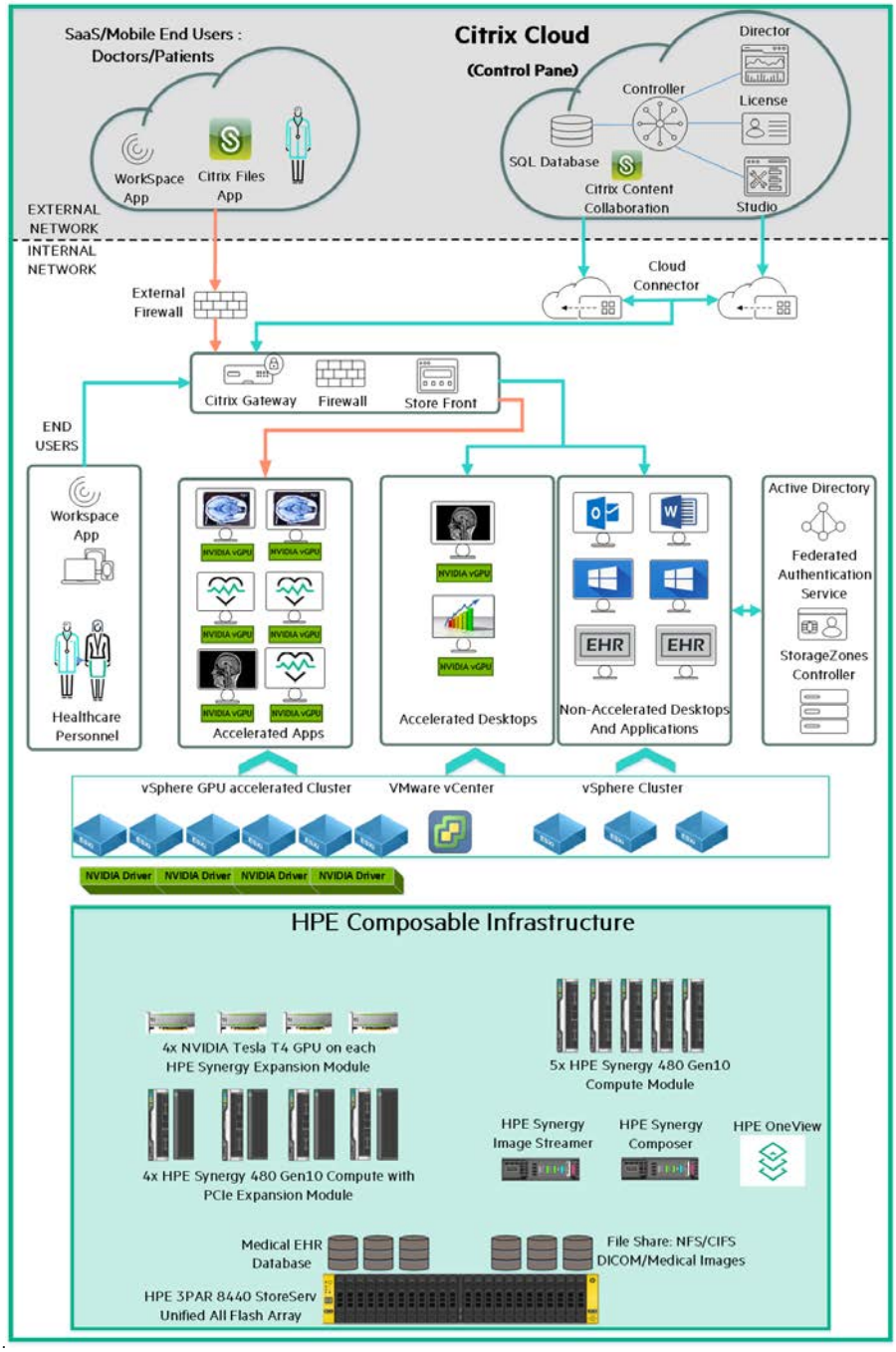

Figure 3. Solution architecture overview

Table 1 summarizes the hardware configuration for this Reference Configuration.

Table 1. List of Hardware components

Hardware	Quantity	Description
HPE Synergy 12000 Frames	3	Infrastructure for compute, storage, fabric and management
HPE Synergy Composer 2	2	Composable Infrastructure management
HPE Synergy 480 Gen10 Compute Module	9	Compute Host
HPE Synergy Image Streamer	2	Infrastructure deployment
HPE 3PAR 8440 AF	1	Storage
HPE PCIe Expansion Module	4	Expansion Module to accommodate NVIDIA GPU
NVIDIA Tesla T4 GPU	12	GPUs dedicated for accelerating virtualized apps and desktops

Table 2 summarizes the software components used for this Reference Configuration.

Table 2. List of Software components

Software	Version	Description
HPE OneView	4.20.01.01	Infrastructure Management Software
Citrix Cloud	--	Cloud Platform for hosting Citrix services for Citrix Content Collaboration and Citrix Virtual Apps and Desktop
Citrix Virtual Apps and Desktops Service	1906	Citrix Cloud offering to deliver virtual apps and desktops from on-premises and/or cloud Resource Locations
Citrix Content Collaboration	5.7	Secure file sharing & sync solution which also allows collaboration.
VMware® vSphere®	6.7	Virtualization and Virtual Machine management platform

Figure 4 shows interaction between Citrix Cloud service, HPE Composable Infrastructure, HPE 3PAR unified Storage hosting the Virtual Desktop Infrastructure and the end users.

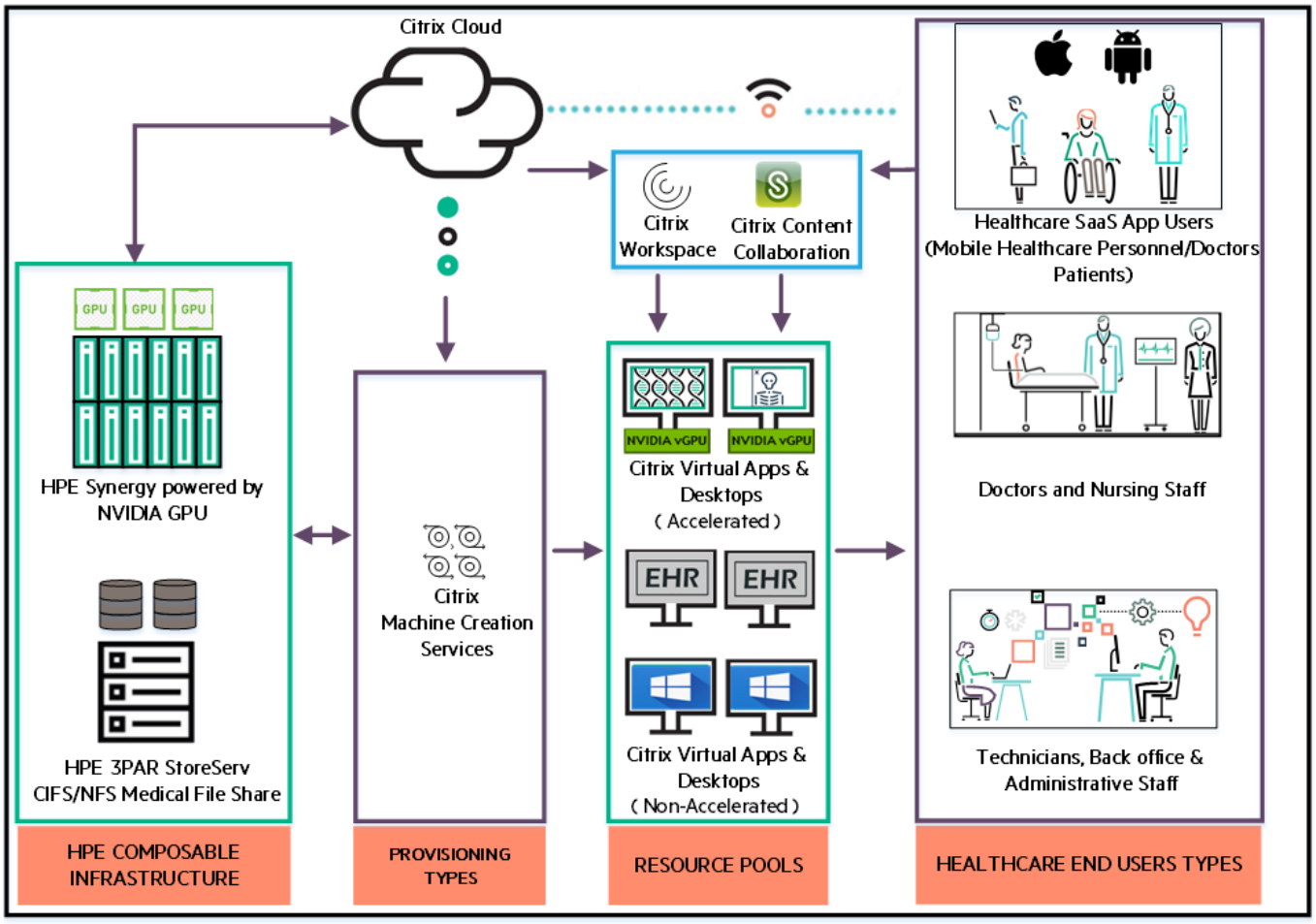

Figure 4. Interaction diagram for this solution

For architecture planning, workers can be categorized into types. The three different user segments and their entitlements are:

- Healthcare users (GPU-Shared) who have a shared NVIDIA GPU infrastructure on HPE Synergy from which accelerated Citrix Virtual Apps are delivered to medical staff such as doctors and nurses.
- Administrative staff, out-patients, and mobile doctors (standard workload) for access to resources that do not typically require graphical acceleration. This user group will be utilizing productivity apps like Microsoft® Office, records management software, and other non-accelerated applications.
- Power users and operators who have a dedicated NVIDIA vGPU profile aligned to the virtual desktop for special use cases like image inferencing software, deep learning and professional 3D applications, etc.

For all three user segments, the access mechanism is via Citrix Workspace app.

Figure 5 highlights three different user segments and their entitlements

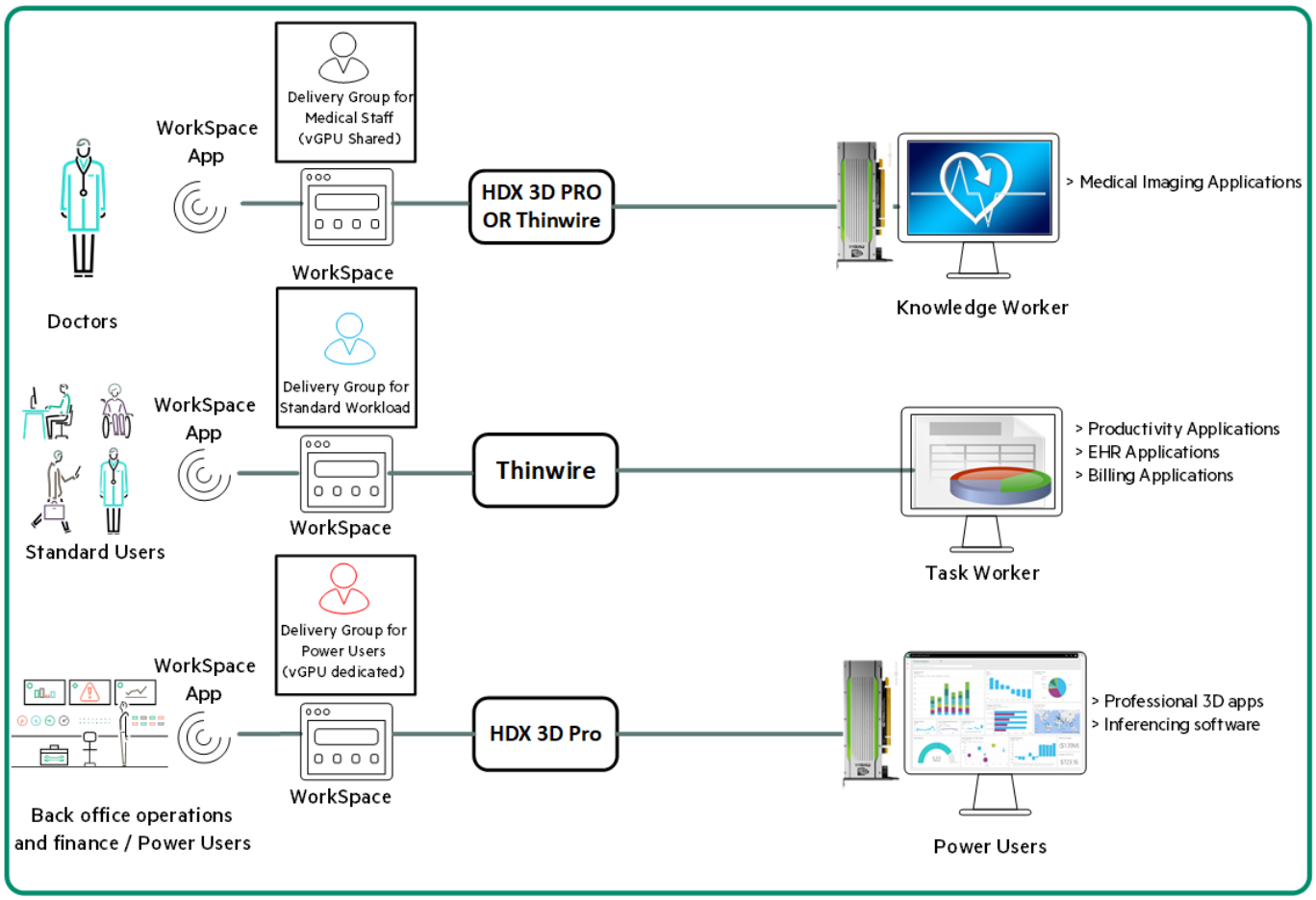

Figure 5. User segments and their workload types

At the core of this solution are the underlying building blocks. These blocks lay the foundation of the digital workspace solution and configured together in different ways to solve unique customer needs. In Figure 6, the Synergy 12000 frame is hosting nine (9) HPE Synergy 480 Gen10 Compute Module and twelve (12) NVIDIA T4 cards for catering the computational and graphical needs of the workload. The HPE 3PAR StoreServ 8440 AF Storage Array provides the speed and storage capacity for our solution.

The key building blocks of the solution are:

- HPE Synergy 480 Gen10 Compute Modules
- HPE PCIe Expansion Module with NVIDIA T4 Cards
- HPE 3PAR StoreServ 8440 AF Storage Array

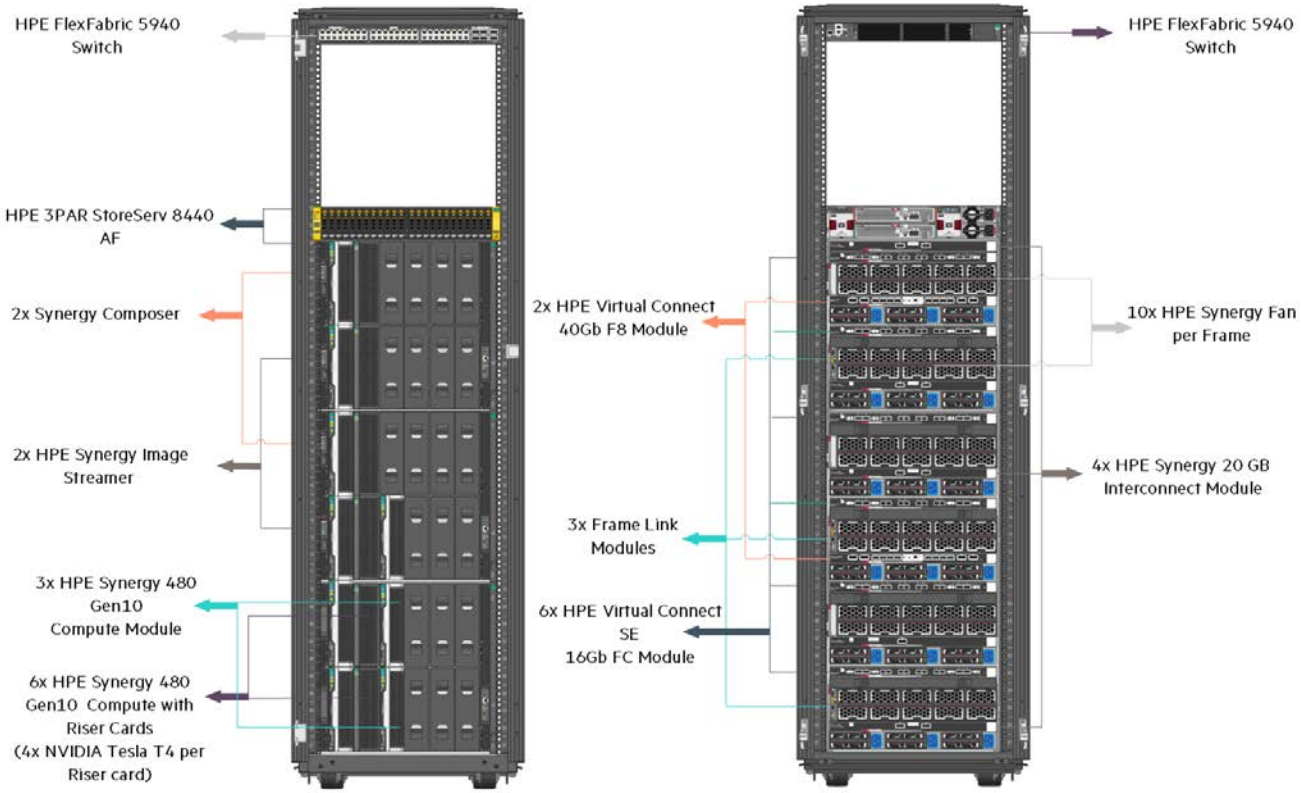

Figure 6. HPE Synergy, HPE 3PAR and NVIDIA components used in this solution

Solution components

HPE Synergy

HPE Synergy, the first platform built from the ground up for composability, offers an experience that empowers IT to create and deliver new value instantly and continuously. A single infrastructure reduces operational complexity for traditional workloads and increases operational velocity for new breeds of applications and services. HPE Synergy empowers IT administrators and developers to use infrastructure as code for

deploying and managing their data center environments. This new approach for composable infrastructure combines true stateless computing with rapid deployment and updates.

Figure 7 highlights the infrastructure evolution from silo to composable.

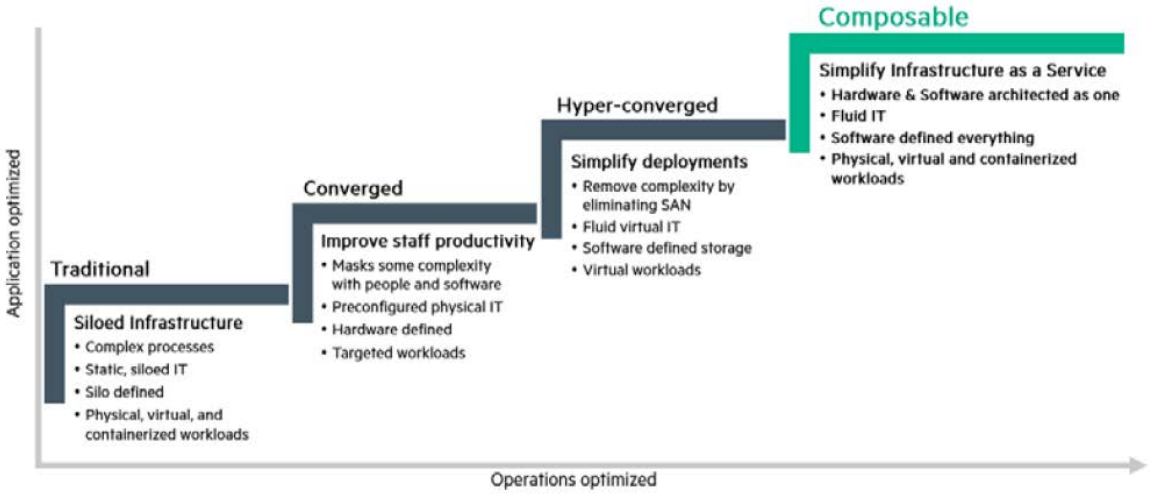

Figure 7. Evolution of infrastructure from traditional silo based systems to modern composable systems

Fluid resource pools

HPE Synergy allows the transformation of traditionally rigid physical systems into flexible virtual resource pools. HPE Synergy creates resource pools of "stateless" compute, storage, and fabric capacity that can be configured almost instantly to rapidly provision infrastructure for a broad range of applications. Now IT can manage both infrastructure growth and shrinkage, from a single fluid pool of resources not stranded in static silos. Smaller infrastructure reduces CapEx, and better management of provisioning reduces power and cooling consumption, minimizing operating expense (OpEx). Fluid resource pools are easily expandable by automatically integrating and assembling additional infrastructure readily composed (and re-composable) to meet the changing workload demands of the business. Figure 8 shows how HPE Synergy's fluid resource pools helps in optimal utilization.

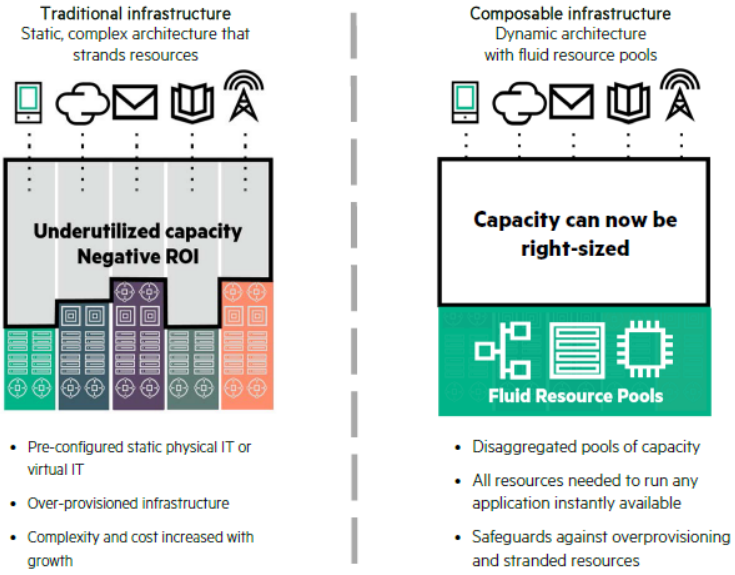

Figure 8. Comparison of traditional versus composable infrastructure

Software-defined intelligence

The software-defined intelligence in HPE Synergy reduces operational complexity and enables IT organizations to make needed programmatic changes quickly and confidently with minimal human intervention. HPE Synergy abstracts operational details and replaces them with high-level, automated operations. HPE Synergy uses templates to enable automatic implementation for change operations such as updating firmware, adding additional storage to a service, or modifying the network. Figure 9 illustrates how HPE Synergy abstracts away operational complexities with help of software-defined intelligence.

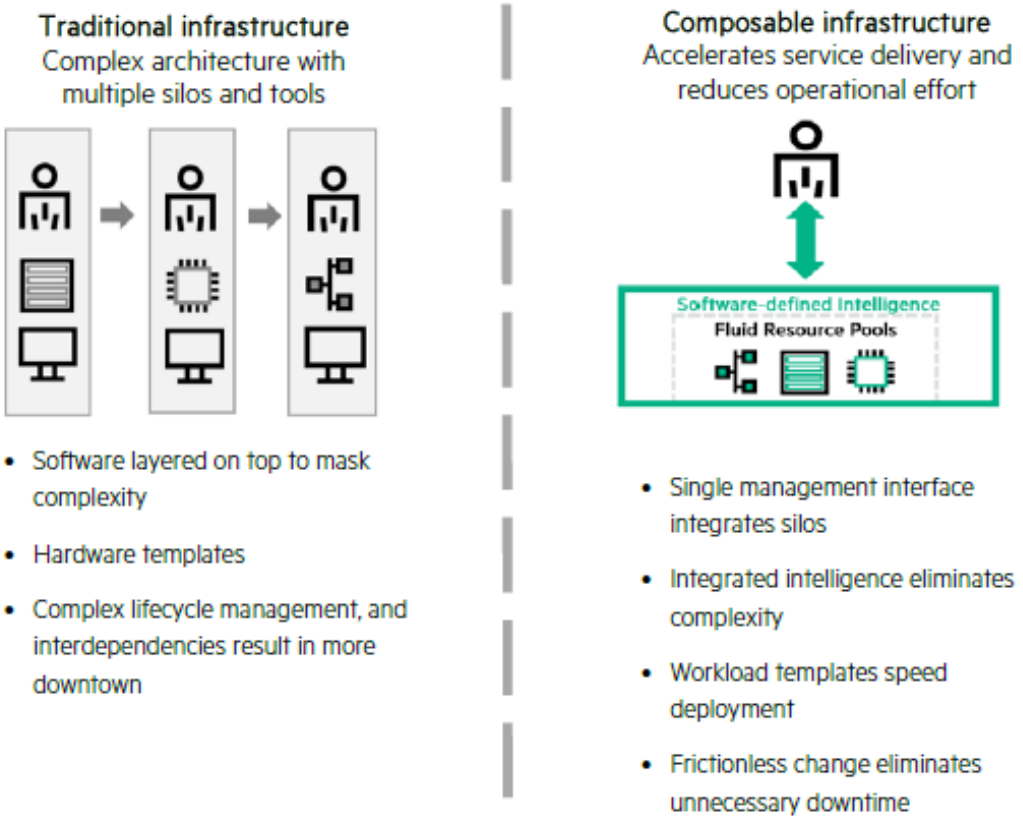

Figure 9. Comparison of traditional versus composable Infrastructure

Unified API

Traditional IT management typically involves low-level API abstractions with a number of orchestration applications. Every device has its own API and each API has error code formats. Using multiple interface requires knowledge of all APIs and which is complex and time-consuming. The task to automate processes across heterogeneous interfaces is complicated since those APIs tend to be at a very low (CLI) level that requires configuration of each component separately. Getting a single server up and running can require as many as 500 or more individual calls to low-level tools to get the infrastructure configured properly. HPE Synergy includes a high-level unified API that brings together all the resources (compute, storage, and fabric) under a single interface with a single data format. Abstracting the API to a high level simplifies programmability.

HPE Synergy 12000 Frame

The HPE Synergy 12000 Frame is a key element of HPE Synergy, providing the base for an intelligent infrastructure with embedded management and scalable links for expansion. The HPE Synergy 12000 Frame is the base infrastructure that pools resources of compute, storage, fabric, cooling, power and scalability. With an embedded management solution combining the HPE Synergy Composer and HPE Synergy Frame Link Modules, IT can manage, assemble and scale resources on demand. The HPE Synergy 12000 Frame is designed for current and upcoming needs with expanded compute and fabric bandwidths.

Figure 10. HPE Synergy 12000 frame

HPE Synergy Composer 2

HPE Synergy Composer 2 provides enterprise-level management to compose and deploy system resources for your applications. This management appliance uses Software-defined intelligence with embedded HPE OneView to aggregate compute, storage and fabric resources in a manner that scales to your application needs, instead of being restricted to the fixed ratios of traditional resource offerings. HPE OneView server profiles and profile templates capture the entire server configuration in one place, enabling administrators to replicate new server profiles and to modify them as needed to reflect changes in the data center. With HPE OneView Rest API and automation tools, the entire process of server personality definition and configuration can be automated. In this solution, the HPE OneView REST API and PowerShell library are used to automate the server profile application to “stateless” servers.

Figure 11. Front view of HPE Synergy Composer node

What’s New?

HPE Synergy Composer 2 comes with Gen10 architecture and equipped with UEFI, High Assurance Boot, iLO5 and two additional 10GB network ports. The memory is increased from 16GB to 64GB and storage is increased from 240GB SATA to 400GB NVMe. The processing power has also been increased from 2-core / 4-threads @ 2.4GHz to 8-core / 16-threads @ 2.0GHz.

HPE Synergy Image Streamer

HPE Synergy Image Streamer is a new approach for administering infrastructure that reduces complexity for traditional workloads and increases operational velocity for the new breed of applications and services. This new approach for composable infrastructure combines true stateless computing with rapid deployment and updates. This management appliance provisions Operating System and applications to work with HPE Synergy Composer for fast Software-defined control over physical compute modules. HPE Synergy Image Streamer enables true stateless computing combined with the capability for image lifecycle management. This management appliance rapidly deploys and updates infrastructure. HPE Synergy Image Streamer adds a powerful dimension to "infrastructure-as-code" – the ability to manage physical servers like virtual machines. In traditional environments, deploying an OS and applications or hypervisor is time consuming because it requires building or copying the software image onto individual servers, possibly requiring multiple reboot cycles. In HPE Synergy, the tight integration of HPE Synergy Image Streamer with HPE Synergy Composer enhances server profiles with images and personalities for true stateless operation.

Figure 12 demonstrates the use of server profiles to deploy software state on a stateless compute module

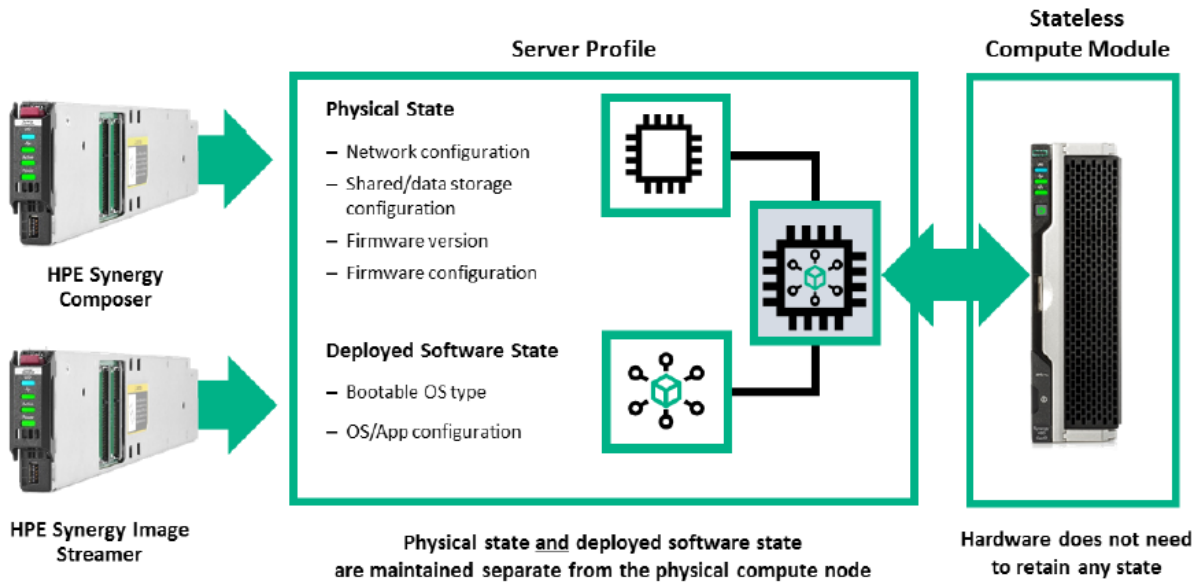

Figure 12. HPE Synergy Composer and HPE Synergy Image Streamer managing compute modules with a server profile

HPE Synergy 480 Gen 10 Compute Module

The HPE Synergy 480 Gen10 Compute Module delivers superior capacity, efficiency, and flexibility in a two-socket, half-height form factor to support demanding workloads. This module provides a composable compute resource that is auto-discovered, inventoried, quickly provisioned, easily managed, and seamlessly redeployed to deliver the right compute capacity for changing workload needs.

Figure 13. Front View of HPE Synergy 480 Gen10 Compute Node

Freed of stranded compute resources, enterprise data centers can now deliver the right compute capacity for changing workload needs. HPE Synergy computes superior, enterprise-grade availability and offers quick and confident infrastructure changes. Change operations such as firmware updates can be applied instantly for initial setup or staged, so they automatically take effect later.

PCIe graphics expansion module

Hewlett Packard Enterprise offers graphics accelerator options for HPE Synergy 480 Gen10 Compute Modules to host multiple graphics cards in MXM format or 2 Full-Length, full height, double wide configuration. The expansion modules are connected to the HPE Synergy 480 Gen10 Compute Modules through PCIe pass-through mezzanine cards located in the compute module. This flexible solution supports the spectrum of graphics cards like NVIDIA Quadro P6000, RTX 6000 and NVIDIA Tesla P40, P6, M10 & T4 to meet the customer requirements by providing high consolidation ratio for accelerated workloads. This solution uses NVIDIA T4 cards with PCIe expansion for accelerating the application and desktop workload.

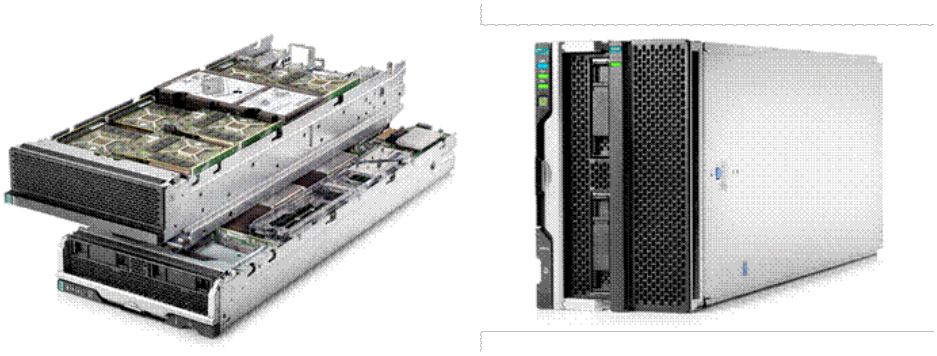

Figure 14. Cross sectional view of a PCIe graphics expansion module

Benefits of HPE Synergy Composable Infrastructure for end-user computing environments

HPE Synergy, a powerful software-defined solution, enables you to compose fluid pools of physical and virtual compute, storage, and fabric resources into any configuration. The benefits are to gain efficiency and control, and deploy IT resources quickly through a single interface.

- HPE Synergy Composer can be leveraged to deliver automation through a unified API that provides a single interface to discover, inventory, configure, provision, update, and diagnose the composable infrastructure in a heterogeneous environment. This fully programmable interface integrates into dozens of popular management tools such as Microsoft System Centre, VMware® vCenter® and DevOps tools such as Chef and Ansible thus making HPE Synergy Composable Infrastructure an ideal platform for Virtual environments. This solution uses simpler and widely used tools, such as Windows PowerShell and VMware PowerCLI to demonstrate the on-demand composability feature of the HPE Synergy platform.
- HPE Synergy's composability features like compose, re-compose, and decompose can be achieved via Unified API in conjunction with VDI desktop pool provisioning using different automation methodologies. Tools and programming languages such as vSphere PowerCLI, Microsoft PowerShell and Python can be used to orchestrate and provision HPE Synergy Fluid Resource Pools comprising of compute, storage and network via Unified API as and when required for VDI Infrastructure. VDI desktop pools can be automated to be provisioned in a timely fashion to consume the readily available composable infrastructure, thus bringing in an efficient and optimized consumption methodology.
- HPE Synergy Composable Infrastructure can be composed and decomposed based on the consumption of desktops by end users in the VDI environment there by optimally consuming the power in the data center. For instance, VMware vSphere Distributed Power Management can be enabled on the vSphere clusters that could trigger ESXi hosts to go into standby mode when there is a decline in consumption of compute resources further enabling the data center administrator to decompose HPE Synergy resources. The result is reduction of power consumption and increase in the green footprint in the data center.
- HPE Synergy Compute Modules in two-slot form factor comprises of expansion module that can host a large number of GPUs when compared to a single form factor compute module. This solution has HPE Synergy with PCIe expansion module, which is capable of hosting 4x NVIDIA Tesla T4 GPUs to address the graphics intensive Virtual Apps and Desktops.
- HPE Synergy Image Streamer appliance provides stateless VMware ESXi hypervisor on HPE Synergy Compute Modules in the composable infrastructure environment. HPE Synergy Image Streamer maintains the hypervisor's golden image and boot volumes in a centralized storage

thereby enabling true stateless computing. In case of replacement of modules due to hardware issue the new nodes can be brought into production rapidly with help of server profiles.

- HPE Synergy Composable Infrastructure can be composed targeting different VDI workloads with respect to GPU and non-GPU workloads. During different time-intervals when the consumption of GPU declines, HPE Synergy Compute Module can be recomposed with a non-GPU server profile for non-accelerated workloads or decompose HPE Synergy resources thereby using the infrastructure optimally.

  Figure 15 demonstrates composable capabilities of HPE Synergy platform.

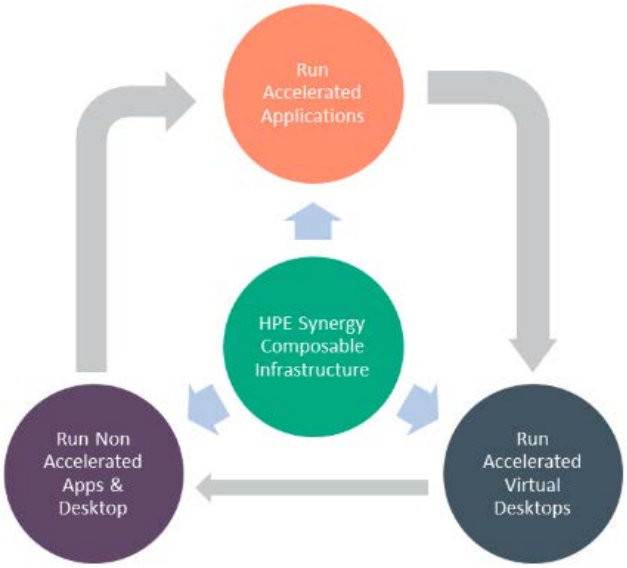

Figure 15. Composable capabilities of HPE Synergy platform

Figure 16 depicts how HPE Synergy includes a high level, unified API that brings together all the resources, compute, storage, and fabric under a single interface.

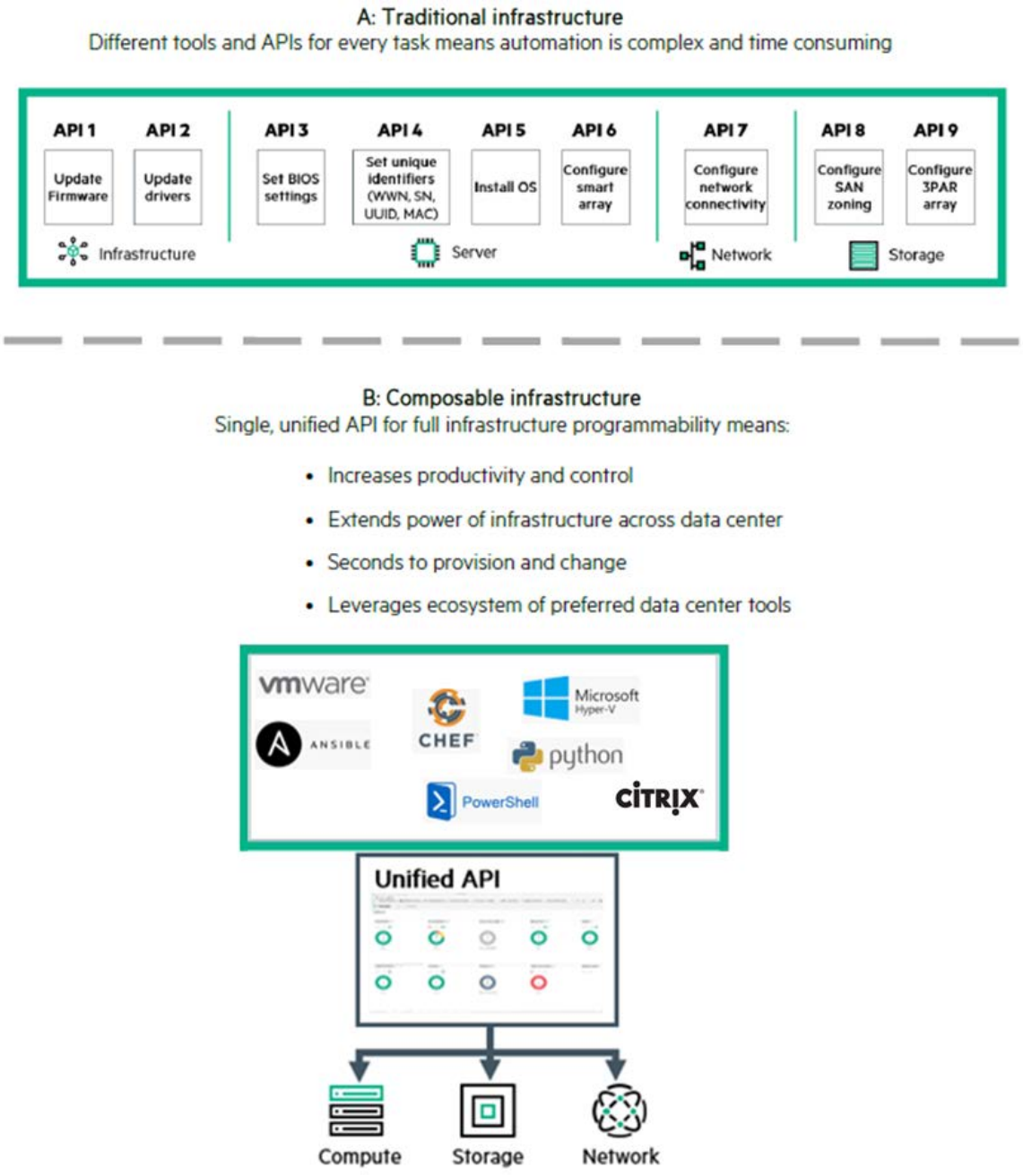

Figure 16. Comparison of programming interface - Traditional versus composable infrastructure

HPE 3PAR 8440 AF

The amount of data generated and processed in today's healthcare environments is massive. The storage solution used should be robust enough to not only support but to accelerate healthcare by enabling organizations to provide personalized care by understanding their needs, reducing risks and expediting decision making. HPE 3PAR StoreServ 8440 Storage solution ensures high throughput and consistently low latency even with demanding workloads of EHR databases, Unstructured DICOM / PACS images along with highly consolidated VDI workload

running medical and productivity applications. Every healthcare organization wants to deliver an exceptional experience and with the advent of software-defined workspaces, and storage plays. One of the most important roles is ensuring that the virtualized desktops are at not only parity but also are faster than their physical counterparts. The HPE 3PAR StoreServ 8440 Storage platform is an enterprise-class flash array that helps you consolidate primary storage workloads (file, block, and object) offering flexible I/O host connectivity without compromising performance, scalability, data services, or resiliency. The purpose of this new HPE 3PAR model based on the proven HPE 3PAR architecture is built for the all-flash consolidation, delivering the performance, simplicity, and agility needed to support your hybrid IT environment. HPE 3PAR StoreServ 8440 Storage is available in a single all-flash model, the 8440, that offers rich Tier-1 data services, quad-node resiliency, fine-grained Quality of Service (QoS), seamless data mobility between systems, high availability through a complete set of persistent technologies, and simple and efficient data protection with a flat backup to HPE StoreOnce backup appliances.

Figure 17. HPE 3PAR StoreServ 8440 Storage system

The HPE 3PAR StoreServ 8440 Storage solution is designed for true convergence of block, file, and object access to help consolidate diverse workloads efficiently. In this Reference Configuration, the EHR database resides on all-flash storage with assured QoS to accelerate EHR applications. Figure 18 depicts how HPE 3PAR file persona provides file storage for high definition medical images on a NFS share via Citrix Content Collaboration to mobilize the health data securely across devices. It is also hosting a CIFS share for storing user profile data to ensure user's personal settings are applied to the user's virtual desktop and applications, regardless of the location and end point device for a seamless experience.

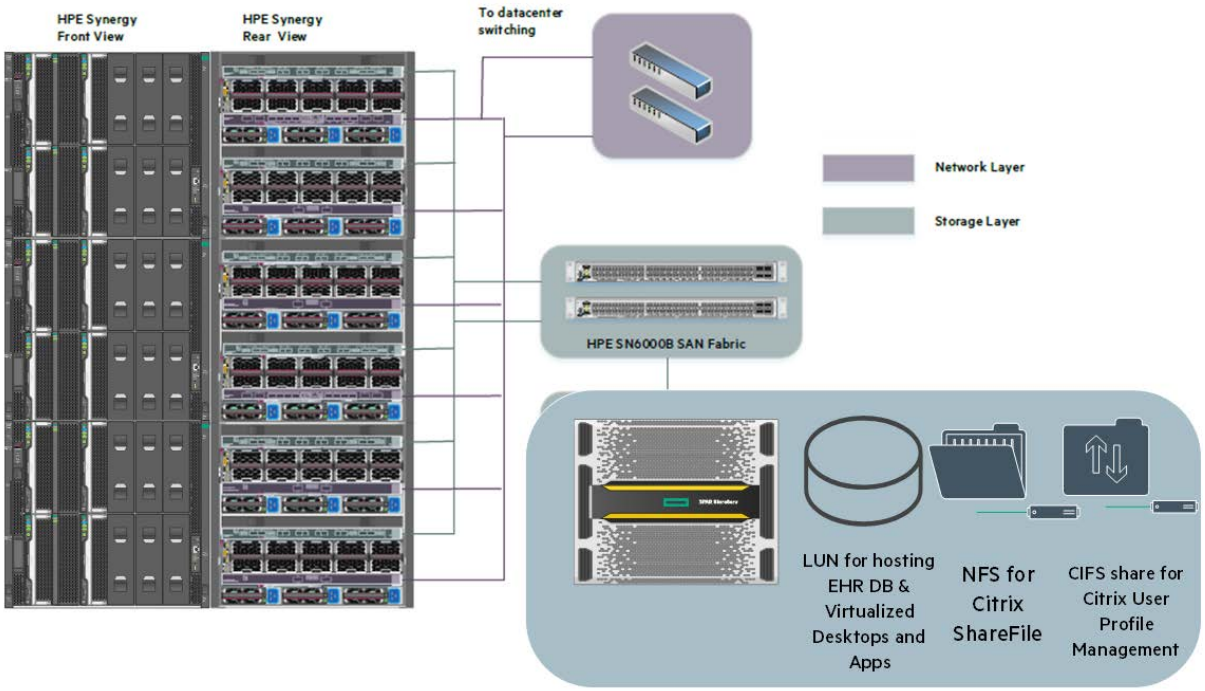

Figure 18. Logical layout of storage solution

NVIDIA GRID

NVIDIA GRID architecture enables NVIDIA GPUs to power Virtual desktops and applications in a data center environment and accelerate workflows resulting in great user experience and improved productivity. Figure 19 depicts how the NVIDIA GRID software works in conjunction with the hypervisor and virtualizes the GPU cards so that they can be shared across multiple virtual desktops or applications, while guest OS on the VM accesses the vGPU like a pass through device.

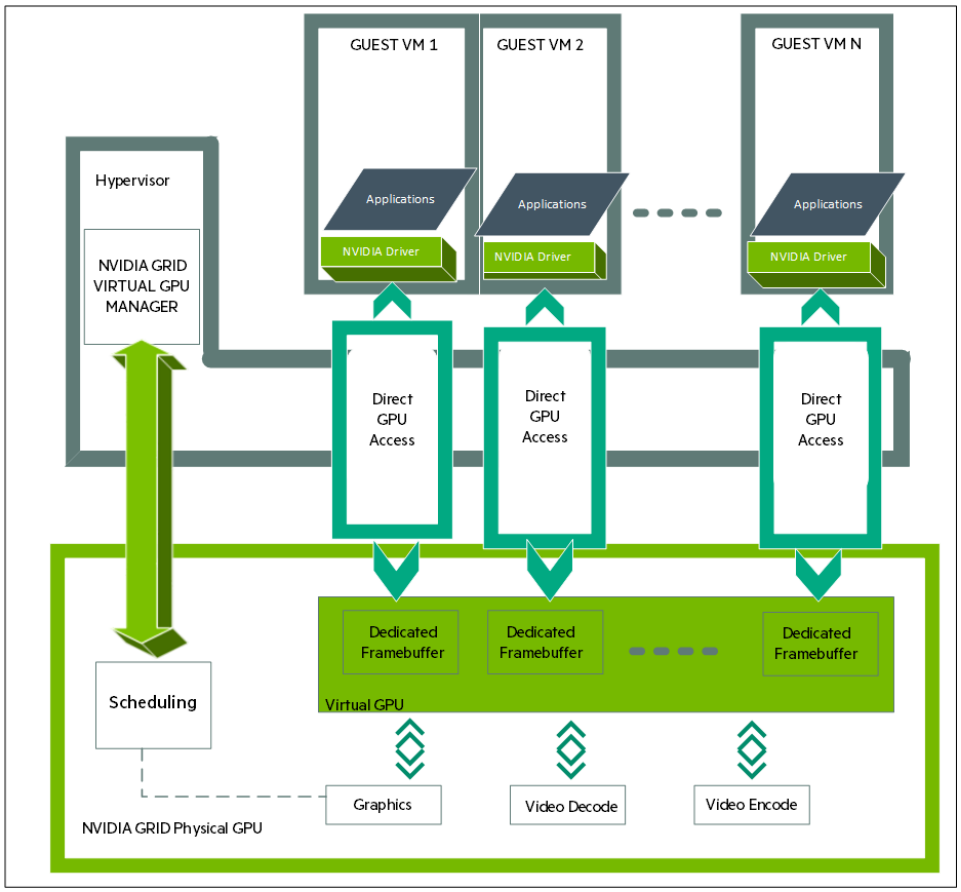

Figure 19. NVIDIA GRID architecture

Table 3 shows the detailed list of NVIDIA GRID components and their function.

Table 3. List of GRID architecture components:

Components	Description
Tesla T4 GPU	NVIDIA Tesla T4 is a graphical processing units based on latest Turing™ architecture for data centers
Virtual GPU Manager	This software runs in the hypervisor layer and allocates physical GPU resources to VMs.
NVIDIA driver	Driver software facilitates direct GPU access to the VMs
Para virtualized interface	Interface for non-performant management operations.
Framebuffer	Frame buffer or graphics memory, is a dedicated resource in NVIDIA GRID for driving video display.
License Server	License manager for NVIDIA GRID.

NVIDIA Licensing

High-end graphics applications and GPU-driven artificial intelligence continue to change the nature of healthcare but at the same time, most of these accelerated applications are used for limited period of time and assigning a dedicated GPU for these type of workload results in underutilized resources and increased licensing and ownership costs. This solution highlights the benefits of accelerating applications rather than the desktops to optimize the utilization of GPU resources, reduced licensing costs and a greener data center. Nevertheless, there are some use cases where the users need high-end graphical computing power throughout their workday and in those cases, assigning dedicated vGPU makes more sense.

Accelerated virtual apps uses GRID vApps licensing model that address the requirements for most of the graphic-intensive workloads in a typical healthcare environment. As an example: viewing DICOM images, running PACS applications and viewing electronic medical records (EMR), viewing and editing very large and complex medical images (PACS), or running a professional 3D application on an accelerated virtual desktop and it can be aligned to the end user who needs a GRID vPC license.

NVIDIA accelerators can be licensed according to the workload as illustrated in Figure 20. GRID vPC, GRID vApps, and Quadro vDWS are available on a per Concurrent User (CCU) model. NVIDIA vGPU editions can be purchased by enterprises either as perpetual licenses or as an annual subscription.

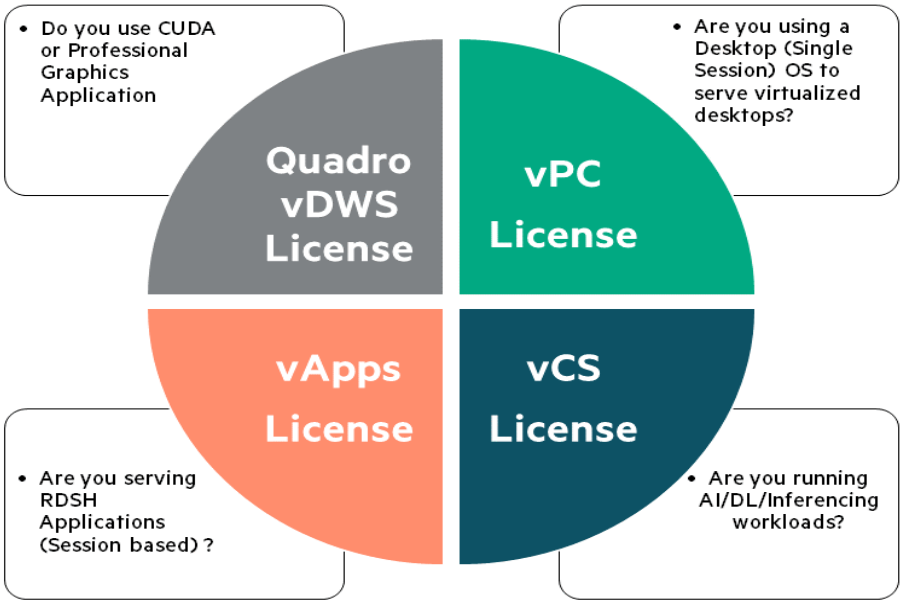

Figure 20. NVIDIA GRID Licensing models and use cases

A perpetual enterprise license allows use of the licensed software indefinitely. Users opt to license using this model are required to subscribe to SUMS for three years. The Support Updates and Maintenance Subscription (SUMS) can be renewed on a yearly basis after the expiration of the initial subscription. Alternatively, NVIDIA GPUs can be licensed with annual enterprise subscription, which is active for a fixed period as defined by the terms of the subscription license. To be kept active, the license will need to be renewed at the end of the subscription period. The subscription license includes the software license and production level SUMS for the duration of the license subscription period.

NVIDIA GRID licensing model supports a broad spectrum of vGPU profiles that helps in aligning resources aptly and avoid under/over provisioning of vGPUs. The servers running on accelerated applications can be configured with 16GB profile, which will be accessed by multiple users. Additionally NVIDIA offers profile sizes of 1GB, 2GB, 4GB, 8GB, and 16GB for Quadro vDWS. Similarly, the graphics workload can be accelerated with 1GB or 2GB profiles for the accelerated desktops.

Note

For latest pricing details, please check https://www.nvidia.com/en-us/data-center/buy-grid/

NVIDIA Tesla T4 GPU

NVIDIA Tesla T4 GPUs are the latest breed of data center based on Turing™ architecture, which enables the acceleration of mixed workloads like virtual machines (VMs) and deep learning, AI, inferencing, high-performance computing (HPC), and other applications. With HPE Composable Infrastructure and NVIDIA GPU dual-mode capability, we can run different workloads.

Table 4 provides information about NVIDIA Tesla T4 specifications.

Table 4. Tesla T4 – Tensor core GPU Specifications

Specifications	
GPU Architecture	NVIDIA Turing
Turing Tensor Cores	320
CUDA Cores	2560
Memory	16GB GDDR6
Memory bandwidth	320 GB/s
Peak single precision floating point performance	8.1 TFlops
Peak double precision floating point performance	254.4 GFlops

NVIDIA Tesla T4 GPUs with their low profile, single slot form factor are extremely powerful and versatile which makes them ideal for data center based deployment and with NVIDIA's Compute Unified Device Architecture (CUDA), the programs get direct access to parallel computing elements of the GPU which allows enterprises to run accelerated workloads.

Figure 21 depicts NVIDIA Tesla T4 GPU.

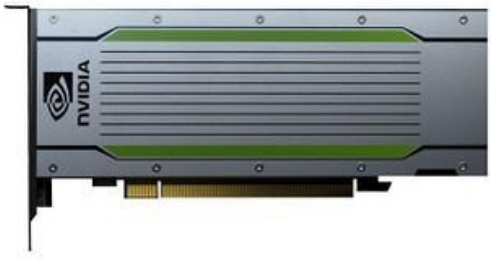

Figure 21. . NVIDIA Tesla T4 GPU

Note

NVIDIA Encode (NVENC) technology with Citrix HDX 3D Pro offloads the encoding of H.265 video output from the server CPU and moves this workload to the GPU resulting in faster processing and improved performance of hosted applications while saving CPU cycles.

Software Components

HPE OneView

HPE OneView is an infrastructure automation engine that simplifies operations to increase the speed of IT delivery for new applications and services. Through software-defined intelligence, HPE OneView brings a new level of automation to infrastructure management by taking a template-driven approach to provisioning, updating, and integrating compute, storage, and networking infrastructure. Designed with a modern, standard-based application-programming interface (API) and supported by a large and growing partner ecosystem. HPE OneView also makes it easy to integrate powerful infrastructure automation into existing IT tools and processes.

Citrix Cloud

Citrix Cloud contains cloud-based services that enables the modern digital workspace with networking and analytics. It's Virtual Apps and Desktops Service helps an organization in offloading the core components of the delivery infrastructure to the cloud, where Citrix manages the installation, maintenance and upgrades of those components while IT can focus on managing applications, desktops, policies and user access.

The Virtual Apps and Desktops Service manages Delivery Controllers, Studio, Director, SQL Servers, the license server, and optionally StoreFront and Gateway. The StoreFront and Gateway is collectively referred to as the Control Layer. The resource, platform, operations, and User Layers are managed within the enterprise, the Access Layer is comprised of the StoreFront, and the Gateway, managed by the customer or Citrix, or you can have those components on-premises.

Figure 22 highlights different Citrix Cloud layers and their components for a better understanding.

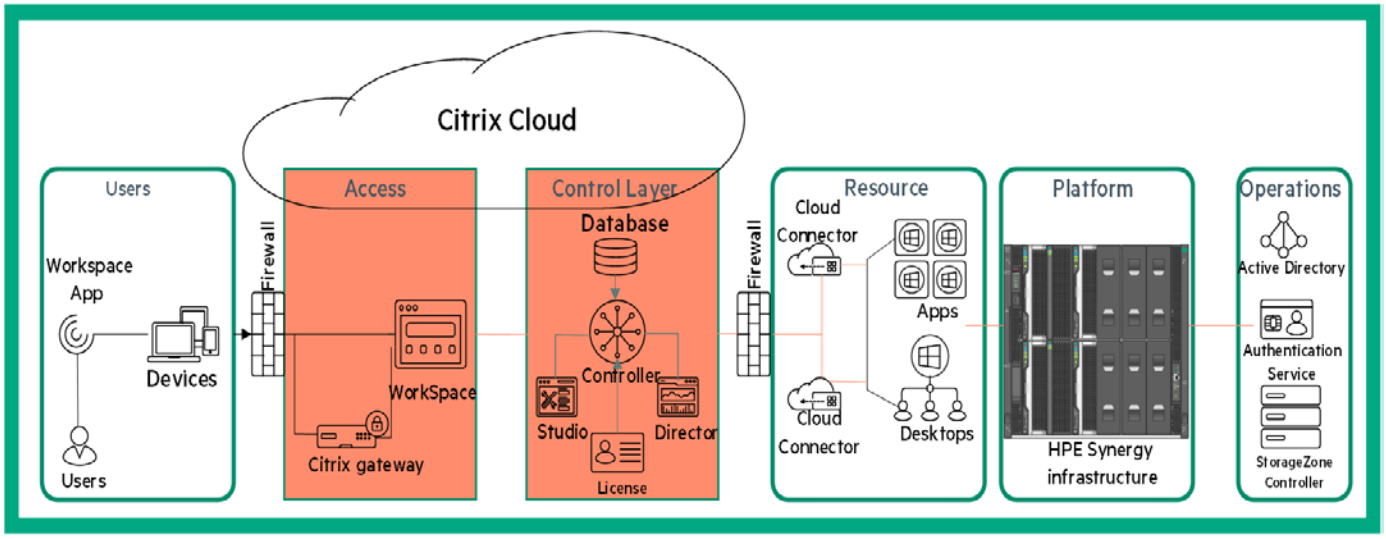

Figure 22. Citrix Cloud layers and their components

These layers are as follows:

- User Layer –This layer encompasses users and their devices.
- Access Layer –This layer defines how a user group gains access to their resources while providing secure access policies and desktop/application stores.
- Resource Layer –This layer is the focal point of the solution as this layer encompasses virtualized apps and desktops, which will be accessed by the end users.
- Control Layer–This layer represents the Citrix management layer supports users accessing resources. The delivery controllers authenticate users and enumerate resources from StoreFront while creating, managing, and maintaining the virtual resources. All configuration information about the Citrix apps and desktop environment is stored within a back end cloud-hosted database and the licensing server manages the licensing.

- With the Citrix Cloud model, you create, deploy, and manage apps and desktop for your users on your preferred platform without the overhead of installation, setup, configuration, maintenance, and monitoring of Citrix Control Layer Infrastructure. By choosing this model, an enterprise can focus on user applications and desktops and offload the deployment, maintenance, and upgrades of core components to Citrix.
- Platform Layer– This layer comprises of infrastructure components like compute, storage and network fabric, hypervisor, virtual environment manager, and this layer hosts resource layer.
- Operations Layer – This layer contains the procedures and tools that support the core product and solution.

Citrix Virtual Apps and Desktop Service

Citrix Virtual Apps and Desktop Service provides a cloud-based management plane, which helps in secure remote delivery of applications and desktop enabling automation and centralized management. The Citrix solution helps end users to access their resources from a plethora of devices like personal PCs, MacBook, thin clients, zero clients, tablets and smartphones.

Citrix Virtual Apps and Desktops addresses multiple challenges faced by healthcare providers:

- Enables secure and quick access to medical applications and data.
- Improves compliance and security.
- Reduces cost and complexity for administrators
- Increase uptime and availability

Citrix Virtual Apps and Desktops manages delivery by utilizing following components.

Table 5. Components of Citrix Virtual apps and Desktop Service

Component	Description
HDX 3D PRO Technology	Remote display protocol that delivers a high-definition 3D user experience of virtual desktops and applications by utilizing graphics processing unit (GPU) for hardware acceleration.
Thinwire	Thinwire is the default display remoting technology used in Citrix Virtual Apps and Desktops.
StoreFront	Facilitates User interface for accessing applications and desktop with self service capabilities.
Controller (Citrix Cloud)	User access management via user or computer based policies.
Studio (Citrix Cloud)	Centralized portal for infrastructure and resource management of application and desktop delivery.
Director (Citrix Cloud)	Monitoring hub for helpdesk staff to get diagnostic information about users, applications and desktop.
Receiver / Workspace	Universal client running on the physical endpoint to access remote applications and desktop.
Licensing Server (Citrix Cloud)	Manages Licensing for Citrix products and services.

Citrix HDX 3D Pro for GPU based medical imaging applications

In this Reference Configuration, Citrix HDX 3D Pro is leveraged to deliver Citrix Virtual Apps and Desktop Service. Citrix HDX 3D Pro is a part of Citrix's HDX remote display technology built on Independent Computing Architecture (ICA) which utilizes graphics processing unit (GPU) for hardware acceleration. This solution uses NVIDIA cards; NVIDIA Encode (NVENC) technology will offload the encoding of H.265 video output from the CPU and move this workload to the GPU resulting in substantially improved graphics for end user and greater overall user density.

Citrix Content Collaboration

Citrix Content Collaboration empowers mobile workstyle and software-defined workspaces by proving a secure file sharing & sync solution to the users and at the same time, it enables management and control to meet corporate data policies and unique compliance requirements. This solution perfectly fits into the healthcare model as it empowers physicians, nurses, and other medical professionals to collaborate effectively and efficiently by providing them a simpler yet secure way to access, share and send files and data while embracing their mobile workstyle.

Citrix Content Collaboration has document workflows and a management module which can be leveraged in a healthcare environment to automate the file management for patients, so that collaborative data and records can be accessed securely on the go by the medical staff.

The access is managed by Citrix files app, which allows secure data sharing and storage along with customizable usage and settings.

Citrix Content Collaboration architecture consists of two distinct regions – the Control plane and StorageZones. The Control plane is hosted in Citrix data centers and is managed by Citrix as a service that is responsible for authentication, access control, brokering and reporting.

StorageZones provides the data storage platforms for Citrix Content Collaboration, which can be customer-managed (On-Premises & AWS/Azure) or Citrix-managed (AWS & Azure). Figure 23 depicts customer-managed StorageZones as this deployment utilizes HPE 3PAR's performance, scalability and data services for reliable and secure storage of critical healthcare data while maintaining sovereignty.

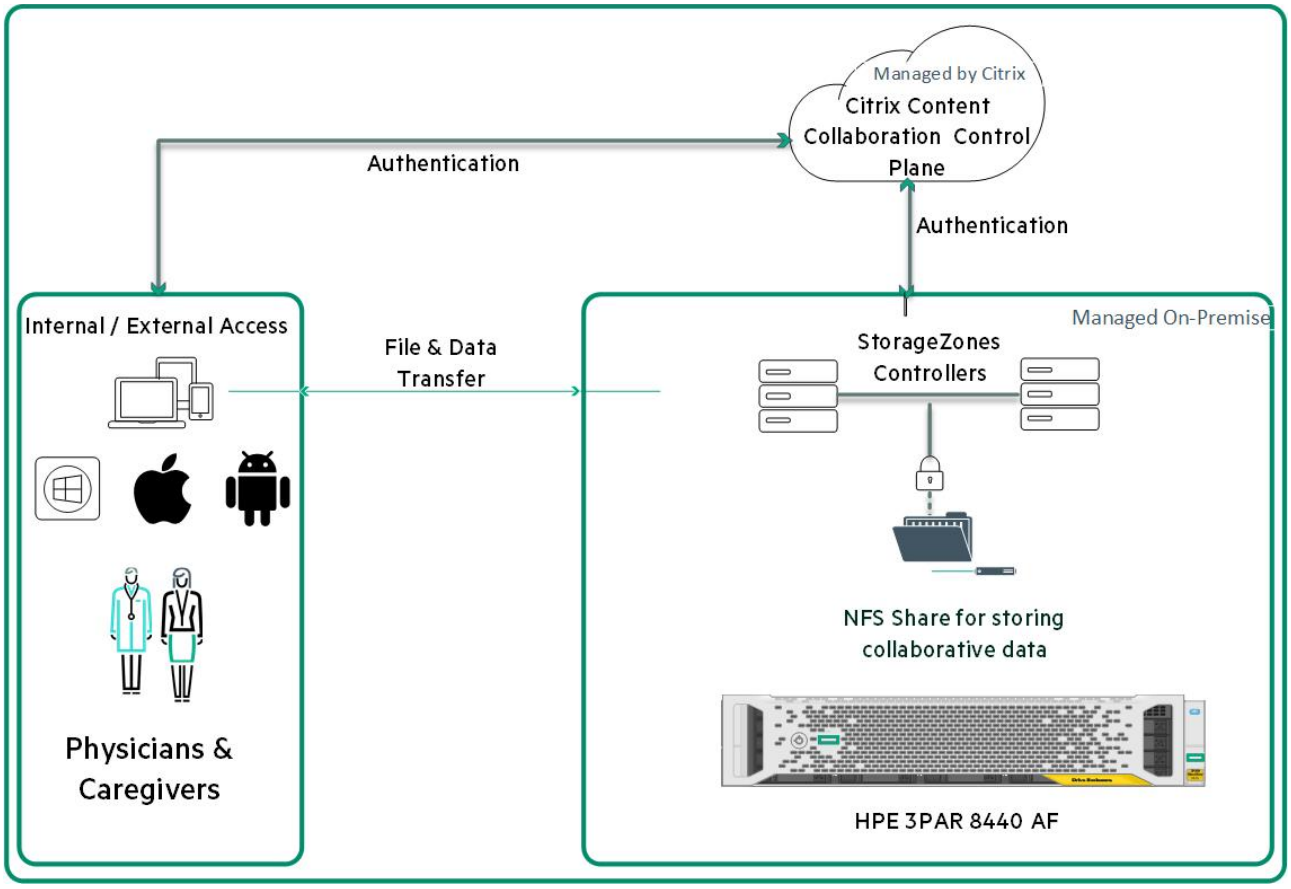

Figure 23. Authentication and data flow diagram of Citrix Content Collaboration

For detailed information about features offered by Content Collaboration visit https://www.sharefile.com/

Summary

Hewlett Packard Enterprise and Citrix have collaborated to develop a joint solution to address the end-user computing infrastructure challenges faced by healthcare organizations, as modern healthcare workspace is undergoing a rapid digital transformation driven by both user's requirements and modern 3D applications. This Reference Configuration provides guidance to healthcare organizations looking for client virtualization to enable a comprehensive, flexible and efficient VDI solution accomplishing requirements of different workloads (accelerated and non-accelerated) and users.

This solution is based on HPE Synergy Composable Infrastructure platform that helps in delivering promised benefits of client virtualization, while overcoming many common challenges.

- Ability to start small and scale out in affordable increments—from pilot to production.
- HPE Synergy 480 Gen10 is a record holder for virtualized performance at 6, 8, 16, and 24 nodes, which translates to unmatched client virtualization performance for a superb end user experience.
- Reduced downtime and maintenance window with true stateless computing.
- High GPU density and freedom to choose multiple GPU types to accelerate the workload.

The key takeaways include:

- VDI implementation is an ideal approach to address compliance, regulations and security requirements, which are imperative to any healthcare organization.
- Empower users with a mobile, BYOD workplace to enable anytime, anywhere, and any-device access to apps and data.
- Improve data security and meet compliance mandates by moving intellectual property and sensitive medical data like EHR & medical reports off the client device and into the data center.
- Ensure business continuity with highly available architecture to reduce downtime for business-critical application.
- Increase the availability of 3D imaging across healthcare organizations using centralized GPU-accelerated applications accessible on demand, which help in effective visualization by physicians.

Note

The solution is capable of delivering pixel-perfect image quality but there are many factors around compliance that need to be considered for medical imaging use cases. This Reference Configuration is not an entire replacement to diagnostics workstations.

Appendix A: HPE Synergy composable infrastructure benefits in VDI/virtual environments

HPE Synergy Composer powered by HPE OneView is core to its composability story. HPE Synergy delivers orchestration via a single interface to discover, inventory, configure, provision, update, and diagnose the composable infrastructure in a heterogeneous environment. This Reference Configuration uses HPE Synergy Composer powered by HPE OneView and HPE Synergy Image Streamer, to demonstrate the on-demand composability feature of the HPE Synergy platform. This Reference Configuration aims to demonstrate the use of fluid resource pools to effortlessly compose and recompose a high-end graphics VDI environment and vice versa, on a single block of disaggregated compute, storage, and fabric infrastructure. It serves as a proof point for dynamically switching workloads from one use case to workloads in an entirely different use case within minutes by using HPE Synergy Composer. Figure 24 shows a HPE Synergy graphical composable flowchart depicting the scenarios and benefits on repurposing infrastructure based on the demand of resources in a real time environment.

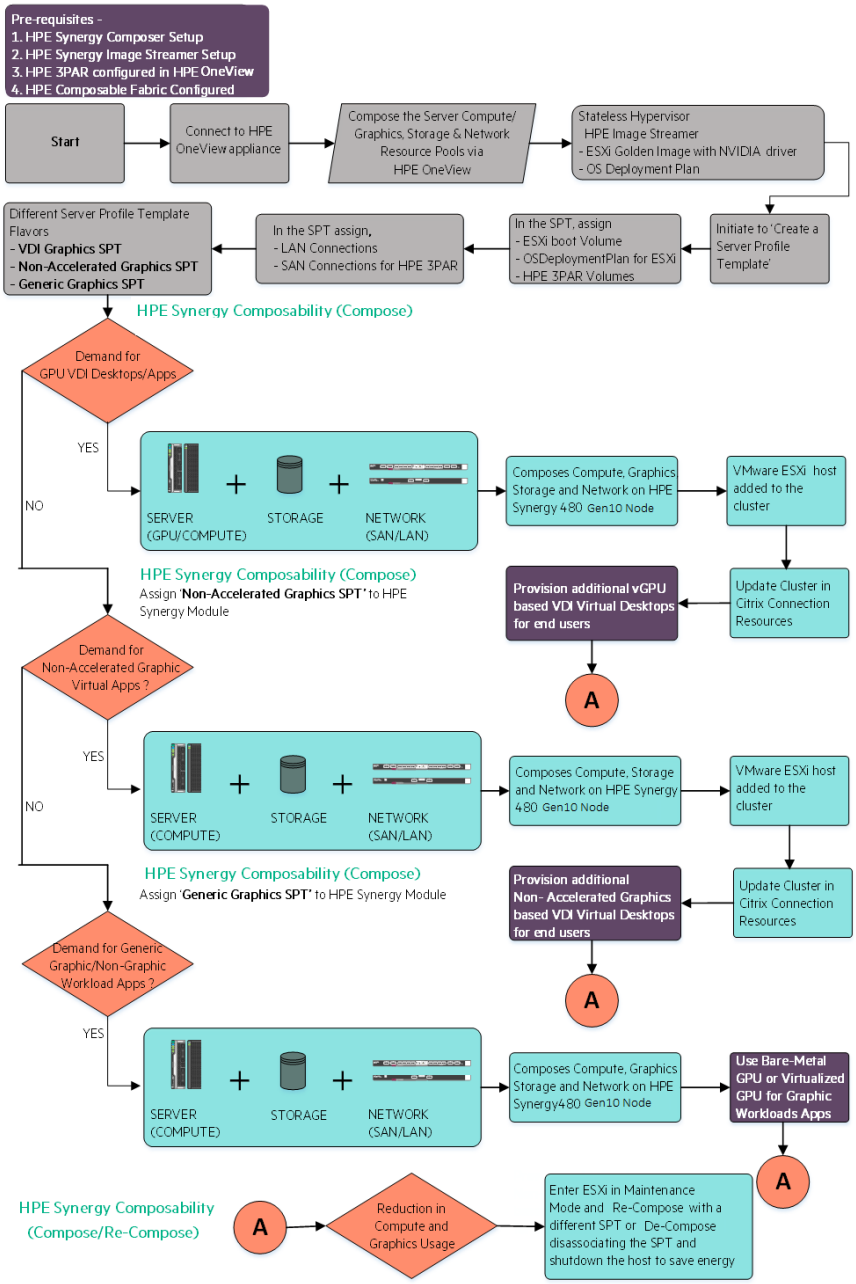

Figure 24. Flowchart showing on-demand composability using HPE OneView

Appendix B: Bill of materials

The following BOMs contain electronic license to use (E-LTU) parts. Electronic software license delivery is now available in most countries. Hewlett Packard Enterprise recommends purchasing electronic products over physical products (when available) for faster delivery and for the convenience of not tracking and managing confidential paper licenses. For more information, please contact your reseller or a Hewlett Packard Enterprise representative.

Note

Part numbers are at time of publication/testing and subject to change. The bill of materials does not include complete support options or other rack and power requirements. If you have questions regarding ordering, please consult with your Hewlett Packard Enterprise reseller or Hewlett Packard Enterprise sales representative for more details. hpe.com/us/en/services/consulting.html.

Table 6 provides Bill of materials for reference purpose.

Table 6. Bill of materials

Qty	Part number	Description
1	P9K10A	HPE 42U 600x1200mm Adv G2 Kit Shock Rack
9	871940-B21	HPE Synergy 480 Gen10 CTO Cmpt Mdl
3	804942-B21	HPE Synergy Frame Link Module
2	804937-B21	HPE Synergy Image Streamer
2	872957-B21	HPE Synergy Composer 2
2	794502-B23	HPE VC SE 40Gb F8 Module
12	K2P89B	HPE 3PAR 8000 1.92TB+SW SFF SSD
1	Q2S13A	HPE 3PAR StoreServ RPS Service Processor
1	H6Y97B	HPE 3PAR 8440 2N+SW Storage Base
4	QK734A	HPE Premier Flex LC/LC OM4 2f 5m Cbl
2	P9Q41A	HPE G2 Basic 4.9kVA/(20) C13 NA/JP PDU
3	797740-B21	HPE Synergy12000 CTO Frame 1xFLM 10x Fan
9	P07351-L21	HPE Synergy 480/660 Gen10 Xeon-G 6254 FIO Kit
9	P07351-B21	HPE Synergy 480/660 Gen10 Xeon-G 6254 Kit
108	P00918-B21	HPE 8GB 1Rx8 PC4-2933Y-R Smart Kit
72	P00924-B21	HPE 32GB 2Rx4 PC4-2933Y-R Smart Kit
6	870753-B21	HPE 300GB SAS 15K SFF SC DS HDD
9	P01367-B21	HPE 96W Smart Storage Battery 260mm Cbl
9	804424-B21	HPE Smart Array P204i-c SR Gen10 Ctrlr
9	777430-B21	HPE Synergy 3820C 10/20Gb CAN
4	779218-B21	HPE Synergy 20Gb Interconnect Link Mod
2	838327-B21	HPE Synergy Y Dual 10GBASE-T QSFP 30m RJ45 XCVR
2	798096-B21	HPE Synergy 12000F 6x 2650W AC Ti FIO PS
3	804938-B21	HPE Synergy 12000 Frame Rack Rail Option

Qty	Part number	Description
1	804943-B21	HPE Synergy 12000 Frame 4x Lift Handle
1	859493-B21	HPE Synergy Multi Frame Master1 FIO
8	804101-B21	HPE Synergy Interconnect Link 3m AOC
2	720199-B21	HPE BLc 40G QSFP+ QSFP+ 3m DAC Cable
6	861412-B21	HPE CAT6A 4ft Cbl
1	JH397A	HPE FF 5940 2-slot Switch
1	JH180A	HPE 5930 24p SFP+ and 2p QSFP+ Mod
1	JH183A	HPE 5930 8-port QSFP+ Module
12	R0W29A	HPE NVIDIA Tesla T4 16GB Computational Accelerator
4	P14255-B21	HPE Synergy 480 Gen10 PCIe x4 Expansion Module

Resources and additional links

HPE Reference Architectures, hpe.com/info/ra

HPE Composable Infrastructure, hpe.com/info/composable

HPE Composable Systems, hpe.com/info/synergy-ra

HPE Synergy, hpe.com/synergy

HPE Synergy Planning Tool, hpe.com/solutions/synergy-planning-tool

HPE 3PAR StoreServ storage, hpe.com/storage/3par

HPE Technology Consulting Services, hpe.com/us/en/services/consulting.html

HPE Healthcare and Life Sciences IT Solutions, https://www.hpe.com/in/en/solutions/healthcare.html

Citrix Virtual Apps and Desktops service, https://docs.citrix.com/en-us/citrix-virtual-apps-desktops-service.html

Citrix Content Collaboration, https://www.citrix.com/en-in/products/citrix-content-collaboration/

Citrix Cloud services, https://docs.citrix.com/en-us/citrix-cloud.html

To help us improve our documents, please provide feedback at hpe.com/contact/feedback.

a50000219enw, October 2019