White paper Cisco public IIIIII CISCO The bridge to possible

Cisco UCS X-Series with Intel Flex Data Center GPUs and VDI

Cisco Intersight Managed Cisco UCS X-Series Modular System with Intel Flex GPU and VDI

Last updated: January, 2024

Contents

Cisco UCS X-Series with the Cisco Intersight platform	3
Cisco UCS X-Series Modular System	5
Why use Intel Flex GPU for graphics deployments on VDI	7
Intel Data Center GPU Flex cards	8
Reference architecture	9
Configure Cisco UCS X-Series Modular System	10
Install Intel GPU software on the VMware ESXi host	14
Configure host graphics settings	15
Intel Data Center Flex GPUs	16
SPECviewperf results with Intel Flex Virtual Workstation profiles	17
Login Enterprise knowledge worker results with Intel Flex 140 profiles	26
Conclusion	29
For more information	29

This document introduces methods for integrating the Cisco UCS[®] X-Series Modular System with Intel[®] Flex GPU 140 and 170 GPU cards with VMware vSphere for Virtual Desktop Infrastructure (VDI) that is ready for the future.

Cisco UCS X-Series with the Cisco Intersight platform

Cisco delivers the future-ready Cisco UCS X-Series Modular System (Figure 1) to support the deployment and management of virtual desktops and applications. Cisco UCS X-Series computing nodes combine the efficiency of a blade server and the flexibility of a rack server. The UCS X-Series is the first system exclusively managed through Cisco Intersight[®] software. It operates your VDI cloud, including complete hardware and software lifecycle management, from a single interface regardless of location. The Cisco UCS X-Series with the Cisco Intersight platform carries forward the main capabilities Cisco has developed through its long history of virtual desktop deployments:

- Availability: VDI must always be on and available anytime and anywhere. Designed for high availability, the UCS X-Series can sustain single points of failure for supporting infrastructure and keep running. The Cisco Intersight platform offers an always-on connection to the Cisco[®] Technical Assistance Center (Cisco TAC), constantly monitoring your environment to help identify configuration or operational issues before they become problems. As the number of users increases, you can easily scale up and add new capabilities, without downtime.
- User productivity: application access fuels user productivity. Support productivity with fast application deployment and upgrades, simplified application and desktop patching and management, and application migration support. Provision desktops instantly when new staff is hired.
- Flexible design: VDI applications range from call-center workers (task users) accessing a few applications, to professionals (knowledge workers) accessing many applications, to power users accessing graphics-intensive applications. The UCS X-Series is flexible enough to house a complete VDI solution within the chassis (depending on storage needs) or as part of a converged infrastructure solution such as FlexPod or FlashStack. We test and document these solutions in Cisco Validated Designs, where every element is fully documented. The system can be purchased or paid for as it is used through Cisco+.
- **Future ready:** Cisco UCS X-Fabric technology is designed to accommodate new devices and interconnects as they become available, including the capability to extend the server's PCIe connectivity to attach additional Graphics Processing Unit (GPU) accelerators for an even more compelling virtual desktop experience.

The Cisco UCS X-Series with the Cisco Intersight platform is future-ready foundational infrastructure that can support your virtual desktops and applications today and for years to come. With Cisco Intersight software, you can simplify management of all your locations: data center, remote office, branch office, industrial, and edge.



Figure 1.

Cisco UCS X-Series Modular System

Cisco UCS X-Series Modular System

The Cisco UCS X-Series includes the Cisco UCS X9508 Chassis, Cisco UCS X210c M7 Compute Node, and Cisco UCS X440p PCIe Node, discussed in this section.

Cisco UCS X9508 Chassis

The Cisco UCS X9508 Chassis (Figure 2) has the following main features:

- This seven-rack-unit (7RU) chassis has eight front-facing flexible slots. These can house a combination of computing nodes and a pool of future I/O resources, which may include GPU accelerators, disk storage, and nonvolatile memory.
- Two Cisco UCS 9108 25G Intelligent Fabric Modules (IFMs) at the top of the chassis connect the chassis to upstream Cisco UCS 6500 Series Fabric Interconnects. Each IFM provides up to 100 Gbps of unified fabric connectivity per computing node.
- Two 100-Gbps to 25-Gbps uplink cables carry unified fabric management traffic to the Cisco Intersight cloud-operations platform, Fibre Channel over Ethernet (FCoE) traffic, and production Ethernet traffic to the fabric interconnects.
- At the bottom are slots ready to house future I/O modules that can flexibly connect the computing modules with I/O devices. Cisco calls this connectivity Cisco UCS X-Fabric technology, with "X" as a variable that can evolve with new technology developments.
- Six 2800-watt (W) Power Supply Units (PSUs) provide 54 volts (V) of power to the chassis with N, N+1, and N+N redundancy. A higher voltage allows efficient power delivery with less copper and reduced power loss.
- Four 100-mm dual counter-rotating fans deliver industry-leading airflow and power efficiency. Optimized thermal algorithms enable different cooling modes to best support the network environment. Cooling is modular so that future enhancements can potentially handle open- or closed-loop liquid cooling to support even higher-power processors.





Figure 2.

Cisco UCS X9508 Chassis, front (left) and back (right)

Cisco UCS X210c M7 Compute Node

The Cisco UCS X210c M7 Compute Node (Figure 3) is the first computing device to integrate into the Cisco UCS X-Series Modular System. Up to eight computing nodes can reside in the 7RU Cisco UCS X9508 Chassis, offering one of the highest densities of computing, I/O, and storage resources per rack unit in the industry.

The Cisco UCS X210c M7 provides these main features:

- **CPU:** the CPU can contain up to two Fourth Generation (4th Gen) Intel[®] Xeon Scalable Processors, with up to 48 cores per processor and 1.5 MB of Level-3 cache per core.
- Memory: the node can house up to thirty-two 256-GB DDR4 4800-megahertz (MHz) DIMMs, for up to 8 TB of main memory. Configuring up to sixteen 512-GB Intel Optane[™] persistent memory DIMMs can yield up to 12 TB of memory.
- **Storage:** the node can include up to six hot-pluggable, Solid-State Disks (SSDs), or Non-Volatile Memory Express (NVMe) 2.5-inch drives with a choice of enterprise-class RAID or passthrough controllers with four lanes each of PCIe Gen 4 connectivity and up to two M.2 SATA drives for flexible boot and local storage capabilities.
- Modular LAN-on-Motherboard (mLOM) Virtual Interface Card (VIC): the Cisco UCS VIC 14425 can occupy the server's mLOM slot, enabling up to 50 Gbps of unified fabric connectivity to each of the chassis IFMs for 100 Gbps of connectivity per server.
- **Optional mezzanine VIC:** the Cisco UCS VIC 15825 can occupy the server's mezzanine slot at the bottom of the chassis. This card's I/O connectors link to Cisco UCS X-Fabric technology that is planned for future I/O expansion. An included bridge card extends this VIC's two 50-Gbps network connections through IFM connectors, bringing the total bandwidth to 100 Gbps per fabric: a total of 200 Gbps per server.
- Security: the server supports an optional Trusted Platform Module (TPM). Additional features include a secure boot Field Programmable Gate Array (FPGA) and Anti-Counterfeit Technology 2 (ACT2) anticounterfeit provisions.



Figure 3.

Cisco UCS X210c M7 Compute Node

Cisco UCS X440p PCIe Node

The Cisco UCS X440p PCIe Node (Figure 4) is the first PCIe resource node to integrate into the Cisco UCS X-Series Modular System. The Cisco UCS X9508 Chassis has eight node slots, up to four of which can be X440p PCIe nodes when paired with a Cisco UCS X210c M7 Compute Node. The X440p PCIe node supports two x16 full-height, full-length dual-slot PCIe cards, or four x8 full-height, full-length single-slot PCIe cards, and requires both Cisco UCS 9416 X-Fabric Modules for PCIe connectivity. This configuration provides up to 16 GPUs per chassis to accelerate your applications with the X440p PCIe nodes. If your application needs even more GPU acceleration, up to two additional GPUs can be added on each X210c M7 compute node. The Cisco UCS X440p supports the following GPU options:

- Intel Flex 140 GPU
- Intel Flex 170 GPU



Figure 4.

Cisco UCS X440p PCIe Node

Why use Intel Flex GPU for graphics deployments on VDI

Intel Data Center GPU Flex Series is easy to deploy, with support for the majority of popular VDI solutions. It delivers software licensing savings with no additional licensing costs for virtualization of GPUs and no management of licensing servers for VDI deployment. GPUs deliver GPU the levels of performance, quality and capability that users expect, with graphics, media, and compute acceleration, as well as CPU utilization offloads.

The Flex Series GPU helps to enable remote access and increase productivity wherever workers are. It offers an efficient, flexible solution to keep employees up and running with the resources, access, and performance they need to be productive and innovative. The use of GPUs also helps make businesses more future-ready, as the foundation for increasingly graphics-intensive office productivity applications used by knowledge workers and as AI becomes more integrated into mainstream workloads.

- Intel Data Center GPU Flex 170 is a 150-watt PCIe card, optimized for performance. It supports up to 16 VMs and is designed for usages where especially high knowledge worker graphics performance is needed.
- Intel Data Center GPU Flex 140 is a smaller, low-profile PCle card at lower power (75 watts) that is ideal for edge and data-center usages. Each PCle card can support up to 12 typical knowledge-worker VDI sessions.

Intel Data Center GPU Flex cards

For desktop virtualization applications, the Intel Data Center GPU Flex 140 and Flex 170 cards are optimal choices for high-performance graphics VDI (Table 1). For additional information about selecting GPUs for virtualized workloads, see the Intel Technical Brief "Intel Flex GPU Positioning."

Table 1.	Intel graphic cards	specifications
----------	---------------------	----------------

	Intel Data Center GPU Flex 140	Intel Data Center GPU Flex 170
Target workloads	Media processing and delivery, Windows a desktop infrastructure, AI visual inference	and Android cloud gaming, virtualized 3
Card form factor	Half-height, half-length, single-wide, passive cooling	Full-height, three-quarters-length, single-wide, passive cooling
Card TDP	75 watts	150 watts
GPUs per card	2	1
GPU microarchitecture	Xe HPG	XE HPG
X cores	16 (8 per GPU)	32
Media engine	4 (2 per GPU)	2
Ray tracing	Yes	Yes
Peak compute (systolic)	8 TFLOPS (FP32/105 TOPS (INT8)	16 TFLOPS (FP32) / 250 TOPS (INT8)
Memory type	GDDR6	GDDR6
Memory capacity	12 GB (6 per GPU)	16 GB
Virtualization instances	SR-IOV (62)	SR-IOV (31)
Operating systems	Windows server and client (10)	Windows server and client (10)
Host bus	PCle Gen 4	PCle Gen 4
Host CPU support	Intel Xeon processors from 3rd Gen up	Intel Xeon processors from 3 rd Gen up

Reference architecture



Figure 5 provides an overview of the environment used in testing.

Figure 5.

Topology diagram

The following hardware components are used:

- Cisco UCS X210c M7 Compute Node (two Intel Xeon Scalable Gold 6448H CPUs at 2.60 GHz) with 1 TB of memory (64 GB x 24 DIMMs at 4800 MHz)
- Cisco UCS X440p PCIe Node
- Cisco UCS VIC 15231 x 100-Gbps mLOM for X-Series compute node
- Cisco UCS PCI mezzanine card for X-Fabric*
- Cisco x440p PCIe Nodes
- Two Cisco UCS 6536 Fabric interconnects
- Intel Flex 140 cards
- Intel Flex 170 cards
- Two Cisco Nexus[®] 93180YC-FX Switches (optional access switches)^{***}

The following software components are used:

- Cisco UCS Firmware Release 5.2(0.230092)
- VMware ESXi 8.0 Update 2a for VDI hosts
- Citrix Virtual Apps and Desktops 2203 LTSR
- Microsoft Windows 10 64-bit
- Microsoft Windows 11 64-bit
- Microsoft Office 2021
- Intel-idcgpu_1.1.0.1286-10EM.800.1.0.20613240_22435138
- Intel-idcgputools_1.1.0.1286-10EM.800.1.0.20613240_22439847
- Intel Graphics Driver: 31.0.101.4830

*This rear-mezzanine card enables connectivity from the X210c compute node to the X440p PCIe node.

Configure Cisco UCS X-Series Modular System

This section describes the Cisco UCS X-Series Modular System configuration.

Hardware configuration

UCSX-GPU-FLEX 140-MEZZ placement

• The Cisco UCS X440p PCIe Node provides two (riser A) or four (riser B) PCIe slots connected to an adjacent Cisco UCS X210c M7 Compute Node to support Intel Flex GPUs (Figures 6 and 7).



Figure 6. Cisco UCS X440p PCIe Node placement

- Riser A: supports up to two dual-width Intel Flex 170 GPUs
- Riser B: supports up to four single-width Intel Flex 140 GPUs



Figure 7.

GPU placement in risers A and B

Note: GPU models cannot be mixed on a server.

Cisco Intersight configuration

The Cisco Intersight platform provides an integrated and intuitive management experience for resources in the traditional data center and at the edge. Getting started with Cisco Intersight is quick and easy.

See Figure 8 for configuring Cisco UCS with Cisco Intersight Managed Mode (IMM).



Figure 8.

Cisco Intersight workflow

The detailed process is available in several guides available at Cisco.com:

- Getting Started with Intersight.
- Deploy Cisco UCS X210c Compute Node with Cisco Intersight Management Mode for VDI.

This environment uses one PCI node with two Intel Flex 170 cards, one PCI node with four Intel Flex 140 cards, and the compute node with two Flex 140 cards (Figures 9, 10, and 11).

eneral Inventory UCS Ser	rver Profile HCL Statistics	Topology	
S Expand All	GPUs		
Motherboard	ID	Slot ID	Model
Boot	PCIe-Node6-GPU1	CHASSIS-SLOT6-RISER1B-SLOT1	UCSC-GPU-FLEX170
Management Controller	PCIe-Node6-GPU3	CHASSIS-SLOT6-RISER2B-SLOT3	UCSC-GPU-FLEX170
∨ CPUs			
Memory			
V Network Adapters			
GPUs			
PCIe-Node6-GPU1			
PCIe-Node6-GPU3			
DCle Deviese			

Figure 9.

Cisco Intersight views of the PCI node inventory with Intel Flex 170

General Inventory UCS Ser	rver Profile HCL Statistics	Topology	
SEXPAND All	GPUs		
Motherboard	ID	Slot ID	Model
Boot	PCIe-Node6-GPU1	CHASSIS-SLOT6-RISER1B-SLOT1	UCSC-GPU-FLEX170
Management Controller	PCIe-Node6-GPU3	CHASSIS-SLOT6-RISER2B-SLOT3	UCSC-GPU-FLEX170
✓ CPUs			
Memory			
V Network Adapters			
∧ GPUs			
PCIe-Node6-GPU1			
PCIe-Node6-GPU3			

Figure 10.

Cisco Intersight views of the PCI node inventory with Intel Flex 140

Expand All	GPUs				
Notherboard	D	Slot ID	Model	PID	Firmware Version
loot	PCle-Node2-GPU1	CHASSIS-SLOT2-RISER1B-SLOT1	UCSC-GPU-FLEX140	UCSC-GPU-FLEX140	DG02_2.2273-6.7
Aanagement Controller	PCIe-Node2-GPU2	CHASSIS-SLOT2-RISER1B-SLOT2	UCSC-GPU-FLEX140	UCSC-GPU-FLEX140	DG02_2.2273-6.7
Pile	PCIe-Node2-GPU3	CHASSIS-SLOT2-RISER2B-SLOT3	UCSC-GPU-FLEX140	UCSC-GPU-FLEX140	DG02_2.2273-6.7
	PCIe-Node2-GPU4	CHASSIS-SLOT2-RISER2B-SLOT4	UCSC-GPU-FLEX140	UCSC-GPU-FLEX140	DG02_2.2273-6.
3PUs PCIe Devices Storage Controllers PM					

Figure 11.

Cisco Intersight views of the compute node inventory with Intel Flex 140

Install Intel GPU software on the VMware ESXi host

This section summarizes the installation process for configuring an ESXi host and virtual machine for GPU support (Figure 12).

- 1. Download the Intel Flex GPU driver pack for VMware vSphere ESXi 8.0 U2.
- 2. Enable the ESXi shell and the Secure Shell (SSH) protocol on the vSphere host from the Troubleshooting Mode Options menu of the vSphere Configuration Console (Figure 13).

Troubleshooting Mode Options	SSH Support
Enable ESXi Shell Disable SSH Modify ESXi Shell and SSH timeouts Modify DCUI idle timeout	SSH is Enabled Change current state of SSH

Figure 12.

VMware ESXi Configuration Console

- 3. Upload the Intel driver (vSphere Installation Bundle [VIB] file) to the /tmp directory on the ESXi host using a tool such as WinSCP. (Shared storage is preferred if you are installing drivers on multiple servers or using the VMware Update Manager.)
- 4. Log in as root to the vSphere console through SSH using a tool such as Putty.
- 5. The ESXi host must be in maintenance mode for you to install the VIB module. To place the host in maintenance mode, use this command:

```
esxcli software component apply -no-sig-check -d /tmp/Intel-idcgpu_1.1.0.1286-
10EM.800.1.0.20613240_22435138.zip
esxcli software component apply -no-sig-check -d /tmp/Intel-idcgputools_1.1.0.1286-
10EM.800.1.0.20613240_22439847.zip
```

6. The command should return output similar to that shown here:

```
[root@vdi-122:/tmp] esxcli software component apply -no-sig-check-d/tmp/Intel-
idcgpu 1.1.0.1286-10EM.800.1.0.20613240 22435138.zip
Installation Result
    Message: The update completed successfully, but the system needs to be rebooted
for the changes to be effective.
    Components Installed: Intel-idcgpu 1.1.0.1286-10EM.800.1.0.20613240
    Components Removed:
    Components Skipped:
      Reboot Required: true
      DPU Results:
[root@vdi-122:/tmp] esxcli software component apply --no-sig-check -d/tmp/Intel-
idcgputools 1.1.0.1286-10EM.800.1.0.20613240 22439847.zip
Installation Result
    Message: The update completed successfully, but the system needs to be rebooted
for the changes to be effective.
    Components Installed: Intel-idcgputools_1.1.0.1286-10EM.800.1.0.20613240
      Components Removed:
     Components Skipped:
     Reboot Required: true
      DPU Results:
[root@vdi-122:/tmp]
```

Exit the ESXi host from maintenance mode and reboot the host by using the vSphere Web Client or by entering the following commands:

```
#esxcli system maintenanceMode set -e false
#reboot
```

INTEL System Management Interface inventory on VMware ESXi compute node with Intel Flex 140

Both cards, Intel Flex 170 and Intel Flex 140, support Error Correcting Code (ECC) memory for improved data integrity.

Configure host graphics settings

After the Intel GPU Software has been installed, configure host graphics in vCenter on all hosts in the resource pool.

1. Select the ESXi host and click the Configure tab. From the list of options at the left, select Graphics under Hardware. Click Edit Host Graphics Settings (Figure 13).

10.10.70.46	Actions 🗸	
Summary Monitor	Configure Permissions VMs Resource Pools Datastores Networks Updates	
Hardware 🗸	GRAPHICS DEVICES HOST GRAPHICS	1
Overview Graphics	Edit Host Graphics Settings	EDIT

Figure 13.

Host graphics settings in vCenter

- 2. Select the following settings (Figure 14):
 - Shared Direct (Vendor shared passthrough graphics)
 - Spread VMs across GPUs (best performance)



Figure 14.

Edit Host Graphics Settings window in vCenter

Host GPU performance during test in vSphere Web Client

Intel Data Center Flex GPUs

Intel's Flex series graphic processors are available in two SKUs: the Flex Series 170 GPU for peak performance and the Flex Series 140 for maximum density. The graphics processor has up to 32 Intel Xe cores and ray tracing units, up to four Intel Xe media engines, AI acceleration with Intel Xe Matrix Extensions (XMXs) and support for hardware-based SR-IOV virtualization. Taking advantage of the Intel oneVPL Deep Link Hyper Encode feature, the Flex Series 140 with its two GPUs can meet the industry's one-second delay requirement while providing 8K 60 real-time transcode. This capability is available for AV1 and HEVC HDR formats.

For more information, see the Virtual GPU Software User Guide.

Note the following configurations:

- SPECviewperf testing: four vCPUs with 16 GB of memory for 8GB Flex 170 GPU
- **SPECviewperf testing:** four vCPUs with 16 GB of memory for 4GB Flex 170 GPU
- SPECviewperf testing: two vCPUs with 4 GB of memory for 3GB Flex 140 GPU
- Login VSI testing: two vCPUs with 4 GB of memory for 2GB

SPECviewperf results with Intel Flex Virtual Workstation profiles

SPECviewperf 2020 is the latest version of the benchmark that measures the 3D graphics performance of systems running under the OpenGL and Direct X APIs. The benchmark's workloads, called viewsets, represent graphics content and behavior from actual applications.

SPECviewperf uses these viewsets:

- 3ds Max (3dsmax-07)
- CATIA (catia-06)
- Creo (creo-03
- Energy (energy-03)
- Maya (maya-06)
- Medical (medical-03)
- Siemens NX (snx-04)
- SolidWorks (solidworks -07)

The following features introduced in SPECviewperf 2020 v3.0 were used for the tests:

- New viewsets are created from API traces of the latest versions of 3ds Max, Catia, Maya, and SolidWorks applications.
- Updated models in the viewsets are based on 3ds Max, Catia, Creo, SolidWorks, and real-world medical applications.
- All viewsets support both 2K and 4K resolution displays.
- User interface improvements include better interrogation and assessment of underlying hardware, clickable thumbnails of screen grabs, and a new results manager.
- The benchmark can be run using command-line options.

SPECviewperf hardware and operating system requirements are as follows:

- Microsoft Windows 10 or Windows 11 or later
- 16 GB or more of system RAM
- 80 GB of available disk space
- Minimum screen resolution of 1920 × 1080 for submissions published on the SPEC website

- OpenGL 4.5 (for catia-06, creo-03, energy-03, maya-06, medical-03, snx-04, and solidworks-07) and DirectX 12 API support (for 3dsmax-07)
- GPU with 4 GB or more dedicated GPU memory
- Flex 140 with 3 GB of frame buffer was configured with 2 Virtual Functions (VFs) per GPU.
- Flex 170 with 8 GB of frame buffer was configured with 2 Virtual Functions (VFs) per GPU.
- Flex 170 with 4 GB of frame buffer was configured with 4 Virtual Functions (VFs) per GPU
 - Example: one Intel Flex 170 GPU is 16 GB, and to get a 4 GB frame buffer profile we had to define 4 VFs (16GB/4 VFs = 4 GB per PCI GPU)

Cisco UCS X210c M7 Compute Node with PCI node (FLEX 140-3GB)

Figures 15 through 19 show the results for Microsoft Windows 10 with four vCPUs and 16 GB of memory and the Intel Flex 140-3GB profile on a Cisco UCS X210c M7 Compute Node with dual Intel Xeon Gold 6448H 2.40-GHz 32-core processors and 2 TB of 4800-MHz RAM.

	SPECviewpe	erf® 2020 Resu l	ts
SCORES		Intel Flex	140 GPU 3GB
CONFIGURATION 3DSMAX-07	SPECviewperf measures the frame ra across a wide variety of applications each composite score below is based	20x1080) te, or frames per second (FPS), at which your _i and usage models. Each viewset represents an on a weighted geometric mean of many differs	graphics card can render scenes application or a usage model, and int scenes and rendering modes.
CREO-03 MAYA-06	Composites by Viewset To explore a viewset in more detail, in comprised, click on the viewset name	ncluding the measured frame rate for each tes	t of which the composite score is
MEDICAL-03	Viewset	Composite Score	Window
SNX-04	3dsmax-07	9.49	1900 x 1060
SOLIDWORKS-07	catia-06	18.27	1900 x 1060
	creo-03	21.93	1900 x 1060
	maya-06	94.63	1900 x 1060
	medical-03	12.13	1900 x 1060
	snx-04	1.53	1900 x 1060
	solidworks-07	31.37	1904 x 1060

Figure 15. SPECviewperf composite scores



Figure 16.

Perfmon virtual machine CPU utilization



Figure 17. Perfmon virtual machine memory utilization



Figure 18.

Perfmon GPU utilization

Cisco UCS X210c M7 Compute Node with PCI node (FLEX 170-8GB)

Figures 20 through 24 show the results for Microsoft Windows 10 with four vCPUs and 16 GB of memory and the Intel Flex 170-8GB profile on a Cisco UCS X210c M7 Compute Node with dual Intel Xeon Gold 6448H 2.40-GHz 32-core processors and 2 TB of 4800-MHz RAM.

	SPECviewp	erf⁰ 2020 Resul	ts
	http:/	//www.spec.org/gwpg	
SCORES		Intel/Elev	470 OCD
CONFIGURATION	Composite Scores (19	20x1080)	
3DSMAX-07	SPECviewperf measures the frame re across a wide variety of applications	ate, or frames per second (FPS), at which your <u>s</u> s and usage models. Each viewset represents an	graphics card can render scenes application or a usage model, and
CATIA-06	each composite score below is based	on a weighted geometric mean of many differe	ent scenes and rendering modes.
CREO-03	Composites by Viewset		
ENERGY-03	To explore a viewset in more detail, i comprised, click on the viewset name	including the measured frame rate for each test e in the menu to the left.	t of which the composite score is
MAYA-06	Viewset	Composite Score	Window
MEDICAL-03	3dsmax-07	47.80	1900 x 1060
SNX-04	catia-06	41.73	1900 x 1060
SOLIDWORKS-07	creo-03	54.31	1900 x 1060
	energy-03		
	energy se	20.37	1900 x 1060
	maya-06	20.37	1900 x 1060 1900 x 1060
	maya-06 medical-03	20.37 227.85 34.97	1900 x 1060 1900 x 1060 1900 x 1060
	maya-06 medical-03 snx-04	20.37 227.85 34.97 1.92	1900 x 1060 1900 x 1060 1900 x 1060 1900 x 1060

Figure 19. SPECviewperf composite scores



Figure 20.

Perfmon virtual machine CPU utilization



Figure 21.

Perfmon virtual machine memory utilization



Figure 22. Perfmon GPU utilization



Figure 23. Perfmon GPU memory utilization

Cisco UCS X210c M7 Compute Node with PCI node (FLEX 170-4GB)

Figures 25 through 29 show the results for Microsoft Windows 10 with four vCPUs and 16 GB of memory and the Intel Flex 170-4GB profile on a Cisco UCS X210c M7 Compute Node with dual Intel Xeon Gold 6448H 2.40-GHz 32-core processors and 1 TB of 4800-MHz RAM.

		Intel Flex 170	GPU with 4GB
SCORES	Composite Scores (19	20x1080)	
ONFIGURATION	SDEOviaunarf mansures the frame re	te or fromer per second (EDS) at which your o	raphics card can randar scapes across
3DSMAX-07	SPECviewperf measures the frame rate, or frames per second (FPS), at which your graphics card can render scenes across a wide variety of applications and usage models. Each viewset represents an application or a usage model, and each		
CATIA-06	composite score below is based on a	weighted geometric mean of many different sce	nes and rendering modes.
CREO-03	Composites by Viewset		
ENERGY-03	To explore a viewset in more detail, i	ncluding the measured frame rate for each test	of which the composite score is
Entereor	comprised, click on the viewset name	e in the menu to the left.	
MAYA-06	Viewset	Composite Score	Window
MAYA-06 MEDICAL-03	Viewset 3dsmax-07	Composite Score 41.27	Window 1900 x 1060
MAYA-06 MEDICAL-03 SNX-04	Viewset 3dsmax-07 catia-06	Composite Score 41.27 42.33	Window 1900 x 1060 1900 x 1060
MAYA-06 MEDICAL-03 SNX-04 SOLIDWORKS-07	Viewset 3dsmax-07 catia-06 creo-03	Composite Score 41.27 42.33 55.54	Window 1900 x 1060 1900 x 1060 1900 x 1060
MAYA-06 MEDICAL-03 SNX-04 SOLIDWORKS-07	Viewset 3dsmax-07 catia-06 creo-03 energy-03	Composite Score 41.27 42.33 55.54 10.65	Window 1900 x 1060 1900 x 1060 1900 x 1060 1900 x 1060
MAYA-06 MEDICAL-03 SNX-04 SOLIDWORKS-07	Viewset 3dsmax-07 catia-06 creo-03 energy-03 maya-06	Composite Score 41.27 42.33 55.54 10.65 242.40	Window 1900 x 1060
MAYA-06 MEDICAL-03 SNX-04 SOLIDWORKS-07	Viewset 3dsmax-07 catia-06 creo-03 energy-03 maya-06 medical-03	Composite Score 41.27 42.33 55.54 10.65 242.40 35.24	Window 1900 x 1060
MAYA-06 MEDICAL-03 SNX-04 SOLIDWORKS-07	Viewset 3dsmax-07 catia-06 creo-03 energy-03 maya-06 medical-03 snx-04	Composite Score 41.27 42.33 55.54 10.65 242.40 35.24 1.95	Window 1900 x 1060 1900 x 1060

Figure 24.

SPECviewperf composite scores







Figure 26.

Perfmon virtual machine memory utilization





Login Enterprise knowledge worker results with Intel Flex 140 profiles

Login Enterprise

Login Enterprise, by Login VSI, is the industry-standard software used to simulate a human-centric workload used for the purpose of benchmarking the capacity and performance of a VDI solutions. The performance testing documented in this Reference Architecture utilized the Login Enterprise benchmarking tool (<u>https://www.loginvsi.com/platform/</u>). The virtual user technology of Login Enterprise simulates real-world users performing real-world tasks while measuring the time required for each interaction. Login Enterprise assesses desktop performance, application performance, and user experience to determine the overall responsiveness of the VDI solution. Using Login Enterprise for VDI capacity planning helps to determine the optimal hardware configuration to support the desired number of users and applications.

About the EUX Score

The Login Enterprise End-User Experience (EUX) Score is a unique measurement that provides an accurate and realistic evaluation of user experience in virtualized or physical desktop environments. The score is based on metrics that represent system resource utilization and application responsiveness. The results are then combined to produce an overall score between 1 and 10 that closely correlates with the real user experience. See the table below for general performance guidelines with respect to the EUX Score.

Score	Assessment
Greater than 8.5	Excellent
7.5-8.5	Very Good
6.5-7.5	Good
5.5-6.5	Fair
Less than 5.5	Poor

See the following article for more information: <u>https://support.loginvsi.com/hc/en-us/articles/4408717958162-</u> Login-Enterprise-EUX-Score-.

Single server with Intel Flex 140 GPU and graphics intensive knowledge worker workloads for a singlesession OS with random sessions with 3 GB virtual function

Figures 30 through 33 show the results for a graphics intensive knowledge worker workload on a Cisco UCS X210c M7 Compute Node with dual Intel Xeon Gold 6448H 2.40-GHz 32-core processors and 1 TB of 4800-MHz RAM running Microsoft Windows 10 64-bit and Office 2021 virtual machines with 4 vCPUs and 8 GB of RAM.

Note: No VMware or Citrix optimizations were applied to the desktop image.



Figure 28. Login Enterprise performance chart



Figure 29. GPU utilization



Figure 30.

Virtual machine processor consumed %



Figure 31. Citrix ICA frame rate (frames per second)

Conclusion

The Cisco UCS X-Series is designed with modularity and workload flexibility. It allows you to manage the different lifecycles of CPU and GPU components easily and independently. The Cisco UCS X210c M7 Compute Node with Intel Xeon CPUs provides a highly capable platform for enterprise end-user computing deployments, and the Cisco UCS X440p PCIe Node allows you to add up to four GPUs to a Cisco UCS X210c Compute Node with Cisco UCS X-Fabric technology. This Cisco architecture simplifies the addition, removal, and upgrading of GPUs to computing nodes, and in conjunction with the Intel Flex Data Center graphics cards, allows end-users to benefit more readily from GPU-accelerated VDI.

For more information

For additional information about topics discussed in this document, see the following resources:

- cisco.com/go/ucsx
- <u>cisco.com/go/intersight</u>
- cisco.com/go/vdi
- Install and configure Intel Flex GPUs on ESXi

Americas Headquarters Cisco Systems, Inc.

San Jose, CA

Asia Pacific Headquarters Cisco Systems (USA) Pte. Ltd. Singapore Europe Headquarters Cisco Systems International BV Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at https://www.cisco.com/go/offices.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: https://www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)

Printed in USA