



テクニカル レポート

VMware Horizon 7 仮想デスクトップ インフラ向け NetApp HCI

パワー ユーザを 3D グラフィックスで支援

ネットアップ Suresh Thoppay
2019 年 8 月 | TR-4792

提携協力 :



概要

このテクニカル レポートは、NVIDIA のグラフィック処理ユニット (GPU) と仮想化ソフトウェアを実装した VMware Horizon 環境で NetApp® HCI 615C を 3D グラフィックスワークフロードに使用するためのガイダンスを提供します。また、SPECviewperf 13 を使用した NetApp HCI 615C の予備テストの結果も記載します。

目次

1 概要	6
2 NetApp HCI	6
2.1 ストレージ ノード	7
2.2 コンピューティング ノード	7
3 NVIDIA ライセンス	8
3.1 GRID vPC	9
3.2 GRID vApps	9
3.3 Quadro vDWS	9
3.4 NVIDIA vCS	9
4 VMware vSphere 上の 3D ワークロード	9
4.1 Virtual Dedicated Graphics	9
4.2 Virtual Shared Graphics	11
5 VMware Horizon	19
5.1 ソリューション リファレンス アーキテクチャ	19
5.2 VMware Horizon Connection Server	19
5.3 Horizon Client	19
5.4 App Volumes	20
5.5 User Environment Manager	20
5.6 SMB ファイル共有	20
5.7 Just-in-Time Management Platform	21
5.8 Universal Access Gateway	21
5.9 VMware vRealize Log Insight	21
5.10 VMware vRealize Operations	21
6 SPECviewperf 13 ベンチマーク	21
6.1 概要	21
6.2 3ds Max (3dsmax-06)	23
6.3 CATIA (catia-05)	27
6.4 Creo (creo-02)	31
6.5 Energy (energy-02)	35
6.6 Maya (maya-05)	39
6.7 Medical (medical-02)	43
6.8 Showcase (showcase-02)	47
6.9 Siemens NX (snx-03)	50
6.10 SOLIDWORKS (sw-04)	54

7 ネットアップ プライベート クラウドとの統合	57
8 NetApp Kubernetes Service との統合	57
9 まとめ	60
付録 A : vSphere の GPU	61
付録 B : T4 CUDA-Z のスクリーンショット	63
付録 C : T4 vGPU の設定	64
詳細情報の入手方法	68
バージョン履歴	69

表一覧

表 1) vDGA – H610C と H615C の比較	11
表 2) T4 vGPU プロファイル	14
表 3) 単一の GPU における NVIDIA T4 vGPU プロファイル	17
表 4) 単一の GPU における vGPU プロファイルの混在	17
表 5) vGPU – H610C と H615C の比較	18
表 6) ハードウェア リスト	22
表 7) ソフトウェア リスト	23

図一覧

図 1) HCI のコンポーネント	7
図 2) H615C 前面	7
図 3) NVIDIA Tesla GPU	8
図 4) VM 上の 3D グラフィックス	12
図 5) NVIDIA vGPU のアーキテクチャ	13
図 6) [Shared Direct] モード	13
図 7) GPU プロファイルの選択	16
図 8) [Group VMs on GPU until full]	18
図 9) ソリューション リファレンス アーキテクチャ	19
図 10) HEVC の有効化	20
図 11) SPECviewperf 13 のビューセット	22
図 12) 3ds Max の総合スコア	24
図 13) 3ds Max : vSphere のコア利用率 – 16Qx1	25
図 14) 3ds Max : vSphere のコア利用率 – 4Qx12	25
図 15) 3ds Max : GPU 利用率 – 16Qx1	26
図 16) 3ds Max : GPU 利用率 – 4Qx12	26
図 17) 3ds Max のサンプル	27

図 18) CATIA の総合スコア	28
図 19) CATIA : vSphere の CPU 利用率 – 16Qx1	29
図 20) CATIA : vSphere の CPU 利用率 – 4Qx12	29
図 21) CATIA : GPU 利用率 – 1x16Q	30
図 22) CATIA : GPU 利用率 – 12x4Q	30
図 23) CATIA のサンプル	31
図 24) Creo の総合スコア	32
図 25) Creo : vSphere の CPU 利用率 – 16Qx1	33
図 26) Creo : vSphere の CPU 利用率 – 4Qx12	33
図 27) Creo : GPU 利用率 – 16Qx1	34
図 28) Creo : GPU 利用率 – 4Qx12	34
図 29) Creo のサンプル	35
図 30) Energy の総合スコア	36
図 31) Energy : vSphere の CPU 利用率 – 16Qx1	37
図 32) Energy : vSphere の CPU 利用率 – 4Qx12	37
図 33) Energy : GPU 利用率 – 16Qx1	38
図 34) Energy : GPU 利用率 – 4Qx12	38
図 35) Energy のサンプル	39
図 36) Maya の総合スコア	40
図 37) Maya : vSphere の CPU 利用率 – 16Qx1	40
図 38) Maya : vSphere の CPU 利用率 – 4Qx12	41
図 39) Maya : GPU 利用率 – 16Qx1	41
図 40) Maya : GPU 利用率 – 4Qx12	42
図 41) Maya のサンプル	42
図 42) Medical の総合スコア	44
図 43) Medical : vSphere の CPU 利用率 – 16Qx1	44
図 44) Medical : vSphere の CPU 利用率 – 4Qx12	45
図 45) Medical : GPU 利用率 – 16Qx1	45
図 46) Medical : GPU 利用率 – 4Qx12	46
図 47) Medical のサンプル	46
図 48) Showcase の総合スコア	47
図 49) Showcase : vSphere の CPU 利用率 – 16Qx1	48
図 50) Showcase : vSphere の CPU 利用率 – 4Qx12	48
図 51) Showcase : GPU 利用率 – 16Qx1	49
図 52) Showcase : GPU 利用率 – 4Qx12	49
図 53) Showcase のサンプル	50
図 54) Siemens NX の総合スコア	51
図 55) Siemens NX : vSphere の CPU 利用率 – 16Qx1	51
図 56) Siemens NX : vSphere の CPU 利用率 – 4Qx12	52

図 57) Siemens NX : GPU 利用率 – 16Qx1	52
図 58) Siemens NX : GPU 利用率 – 4Qx12	53
図 59) Siemens NX のサンプル	53
図 60) SOLIDWORKS の総合スコア	54
図 61) SOLIDWORKS : vSphere の CPU 利用率 – 16Qx1	55
図 62) SOLIDWORKS : vSphere の CPU 利用率 – 4Qx12	55
図 63) SOLIDWORKS : GPU 利用率 – 16Qx1	56
図 64) SOLIDWORKS : GPU 利用率 – 4Qx12	56
図 65) SOLIDWORKS のサンプル	57
図 66) ハイブリッド クラウド	58
図 67) Hybrid Cloud Control	58
図 68) ネットアップ クラウド サービスの有効化	59
図 69) Bitfusion FlexDirect	60
図 70) vSphere 6.7 Update 1 の T4	61
図 71) vSphere 6.7 Update 2 の T4	61
図 72) NVIDIA vGPU を使用した T4	62
図 73) CUDA-Z のスクリーンショット	63

1 概要

仮想デスクトップとホスト型アプリケーションを使用すると、スマートフォンやタブレットなどのあらゆるデバイスから安全に情報にアクセスできます。ネットアップとネットアップ パートナーは、多くのお客様に対して Virtual Desktop Infrastructure (VDI; 仮想デスクトップ インフラ) 環境の実装を支援してきました。ハードウェア グラフィックス アクセラレーションを求める声は大きくなっています。NetApp HCI のポートフォリオもさまざまなビジネス ニーズに対応するために拡張されました。本ドキュメントでは、AutoDesk、Dassault Systemes、Siemens などのさまざまな製品を対象とした、仮想化 3D グラフィックス ワークロード向けのソリューションについて説明します。

NetApp HCI には、Intel の拡張性に優れた第 2 世代プロセッサや各種メモリ構成などのコンポーネントが搭載されています。コンピューティング ワークロードの実行には、VDI に加えて、拡張性を重視する場合は NVIDIA M10 GPU を、柔軟性を重視する場合は NVIDIA T4 GPU を使用することもできます。ストレージ リソースのニーズに対しては、シンプロビジョニング、圧縮、重複排除などの効率化機能を備えた iSCSI ストレージで応えます。NetApp HCI には、Mellanox CX-4 デュアルポート 2x25GbE ネットワーク コントローラと 1GbE ベースボード管理コントローラがそれぞれ 1 台含まれています。

VMware Horizon 7 には、仮想デスクトップやホスト型アプリケーションのニーズに対応するさまざまな実装オプションが用意されています。セッション共有機能では最大 10 人のユーザによるコラボレーションが可能で、ビデオ コンテンツも問題なく共有できます。

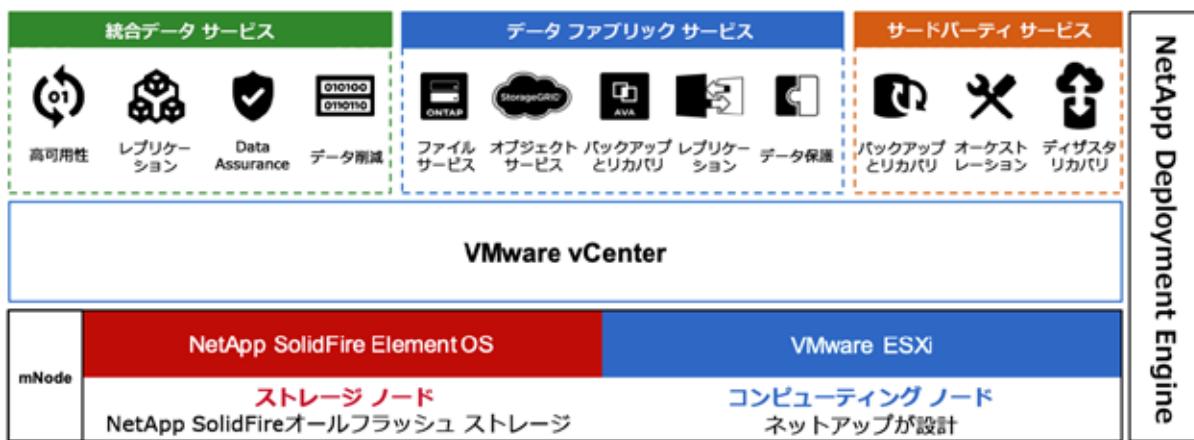
NetApp HCI を使用すると、仮想化 3D グラフィックス ワークロードの課題を克服できるだけでなく、NetApp Kubernetes Service と NVIDIA GPU Cloud を使用すればディープ ラーニング環境も迅速に構築することが可能となります。さらにネットアップ データ ファブリックを使用すると、NetApp HCI のメリットをより強化できます。

2 NetApp HCI

NetApp HCI には、ストレージ ノードとコンピューティング ノードの両方が含まれます。モデルに応じて、2 ラック ユニットまたはシングルラック ユニットのどちらかのシャーシを使用します。仮想マシン (VM) の導入に必要なインストールと設定は、NetApp Deployment Engine (NDE) で自動化されています。コンピューティング クラスタは VMware vCenter で、ストレージ クラスタは NDE で導入された vCenter Plug-in で管理します。管理 VM (mNode) は、NDE の一部として導入されます。mNode は次の処理を担当します。

- バージョンのアップグレード
- vCenter へのイベントのプッシュ
- vCenter Plug-in の管理
- サポート用の VPN トンネル
- NetApp Active IQ コレクタ
- ネットアップ クラウド サービスをオンプレミスに拡張し、ハイブリッド クラウド インフラを実現

図 1) HCI のコンポーネント



2.1 ストレージ ノード

ストレージ ノードには、半幅または全幅サイズのラック ユニットを使用できます。半幅のラック ユニットは、ストレージ ノードとコンピューティング ノードのどちらかを格納できる 2 ラック ユニット シャーシに配置されます。ストレージ ノードは最低 4 つ必要で、シャーシは最大 40 ノードまで拡張可能です。ストレージ クラスタは、複数のコンピューティング クラスタ間で共有できます。すべてのストレージ ノードには、書き込みパフォーマンスを向上させるためにキャッシュ コントローラが搭載されています。1 つのノードで、4K ブロック サイズで 5 万または 10 万 IOPS を実現します。

NetApp HCI ストレージ ノードでは NetApp Element[®] ソフトウェアが実行され、サービス品質 (QoS) 機能を使用して最小、最大、バーストの制限を指定できます。ストレージ クラスタにはタイプの異なるストレージ ノードを混在させることができますが、1 つのストレージ ノードの容量は合計容量の 1/3 以下にする必要があります。

2.2 コンピューティング ノード

コンピューティング ノードには、半幅、全幅、2 ラックサイズのラック ユニットを使用できます。H410C と H610C には Intel スケーラブル プロセッサ (Skylake) が、H615C には Intel 第 2 世代 スケーラブル プロセッサ (Cascade Lake) が搭載されています。GPU を搭載したコンピューティング モデルは 2 つあり、H610C は NVIDIA M10 カードを 2 基、H615C は NVIDIA T4 カードを 3 基搭載しています。

図 2) H615C 前面



NVIDIA T4 には RT コアが 40 基搭載されており、リアルタイム レイ レーシングに必要なコンピューティングを提供します。デザイナーやエンジニアと同じサーバ モデルをアーティストが使用して、水面に反射する光を現実のように表現したリアルな画像を作成できるようになりました。この RTX 対応 GPU は、毎秒最大 5 ギガ レイのリアルタイム レイ レーシング パフォーマンスを実現します。NVIDIA T4 を Quadro Virtual Data Center Workstation (Quadro vDWS) ソフトウェアと組み合わせて使用することで、アーティストは影、反射、屈折を正確に再現した、写真のようにリアルなデザインをどこからでもあらゆるデバイス上に作成できます。

Tensor コアは、ディープ ラーニング推論ワークロードの実行を可能にします。Quadro vDWS を搭載した T4 でディープ ラーニング推論ワークロードを実行した場合の速度は、CPU のみのサーバで動作する VM の最大 25 倍です。1 ラック ユニットに NVIDIA T4 カードを 3 基搭載した H615C は、グラフィックスとコンピューティングの負荷の高いワークロードに最適なソリューションです。

図 3 は、NVIDIA GPU カードと機能比較の一覧です。

図 3) NVIDIA Tesla GPU

仮想化に推奨されるNVIDIA GPU							
	V100	RTX 8000	RTX 6000	P40	T4	M10	P6
GPU	NVIDIA Volta x1	NVIDIA Turing x1	NVIDIA Turing x1	NVIDIA Pascal x1	NVIDIA Turing x1	NVIDIA Maxwell x4	NVIDIA Pascal x1
CUDAコア数	5,120	4,408	4,608	3,840	2,560	2,560 (GPUあたり640)	2,048
Tensorコア数	640	576	576	—	320	—	—
RTコア数	—	72	72	—	40	—	—
QoS保護 (GPU スケジューラ)	✓	✓	✓	✓	✓	—	✓
タイプ マイグレーション	✓	✓	✓	✓	✓	✓	✓
マルチvGPU	✓	✓	✓	✓	✓	✓	✓
メモリ サイズ	32 / 16GB HBM2	48GB GDDR6	24GB GDDR6	24GB GDDR5	16GB GDDR6	32GB GDDR5 (GPUあたり8GB)	16GB GDDR5
vGPUプロファイル	1GB, 2GB, 4GB, 8GB, 16GB, 32GB	108, 208, 308, 408, 608, 808, 12GB, 16GB, 32GB	108, 208, 308, 408, 608, 808, 12GB, 24GB	108, 208, 308, 408, 608, 808, 12GB, 24GB	108, 208, 408, 8GB, 16GB	0.5GB, 1GB, 2GB, 4GB, 8GB	108, 208, 408, 8GB, 16GB
フォーム フラグ	PCIe 3.0デュアルスロットおよびSXM2 (ラック サー/)	PCIe 3.0デュアルスロット	PCIe 3.0デュアルスロット	PCIe 3.0デュアルスロット (ラック サー/)	PCIe 3.0シングルスロット (ラック サー/)	PCIe 3.0デュアルスロット (ラック サー/)	MXM (ブレード サー/)
電源	250W / 300W (500W)	295W	295W	250W	70W	225W	90W
サーマル	パッシブ	アクティブ	アクティブ	パッシブ	パッシブ	パッシブ	ペア ポード
ユースケース	Quadro vDWSを使用する マルチ ハイエンドレンダリング、 シミュレーション、 3Dデザイン、P100に 最高なアップグレード	Quadro vDWSを使用する ハイエンドレンダリング、 3Dデザイン、 クリエイティブ ワークフロー	Quadro vDWSを使用する ハイエンドレンダリング、 3Dデザイン、 クリエイティブ ワークフロー	Quadro vDWSを使用する ハイエンドレンダリング、 3Dデザイン、 クリエイティブ ワークフロー	Quadro vDWSを使用する エンターテイメント/エンターテイメント の3Dデザイン、エンジニアリング、3Dデザイン、 クリエイティブ ワークフロー	最高レベルの密度と拡張性 コスト (TCO)、マルチモニタサポートが必要と する、最先端のビジュアル NVIDIA GRIDソフトウェアを使用するフレッジ ワーク、一括の高効率、 低消費電力のGPU アップグレード	フレッジ ワーク、 NVIDIA GRID vPC / vAppsを使用
							フレッジ ワーク フォーム フレッジのGPUが必要と するお客様向け：MSに 最適なアップグレードバス

M10 GPU は、現在でもナレッジワーカーのユースケースに最適な TCO ソリューションです。ただし、さまざまなユースケースに使用できる GPU での標準化を希望する場合には、T4 が最良の代替ソリューションです。ユースケースの例としては、仮想ワークステーションを使用した、グラフィックスパフォーマンス、リアルタイムのインタラクティブ レンダリング、推論の向上が挙げられます。T4 では、同じ GPU リソースを使用して異なるワークロードを実行できます。たとえば、日中は VDI を実行し、同じリソースを使用して夜間にコンピューティング ワークロードを実行できます。

コンピューティング ノード H610C が 2 ラック ユニットであるのに対して H615C は 1 ラック ユニットのため、消費電力も少なくなります。H615C は、H.264 および H.265 (High Efficiency Video Coding [HEVC]) 4:4:4 のエンコードとデコードをサポートします。また、VP9 デコーダもサポートします。YouTube が提供する WebM コンテナ パッケージでもビデオに VP9 コーデックを使用しているため、VP9 デコーダはより一層主流となりつつあります。

コンピューティング クラスタ内のノード数は VMware によって決まり、現在は 64 個です。 Enhanced vMotion Compatibility (EVC) が有効な場合、クラスタ内に異なるモデルのコンピューティング ノードを混在させることができます。デフォルトのグラフィックス設定 (Virtual Shared Graphics モード) を使用する GPU ノードの場合、クラスタ内にコンピューティング モデルを混在させることができます。

3 NVIDIA ライセンス

NetApp HCI H610C または H615C を使用している場合、GPU のライセンスはライセンスの再販権を持つ NVIDIA パートナーから購入する必要があります。NVIDIA パートナーは [パートナー検索サイト](#) で検索できます。vGPU や Tesla などで検索してください。

NVIDIA vGPU ソフトウェアには、次の 4 つのエディションがあります。

- NVIDIA GRID Virtual PC (GRID vPC)
- GRID Virtual Applications (GRID vApps)
- NVIDIA Quadro vDWS
- NVIDIA Virtual ComputeServer (NVIDIA vCS)

3.1 GRID vPC

GRID vPC は、PC の Windows アプリケーション、ブラウザ、および高解像度ビデオに優れたユーザ エクスペリエンスを提供する仮想デスクトップを必要とするユーザに最適な製品です。NVIDIA GRID Virtual PC は仮想環境のユーザにネイティブ エクスペリエンスを提供し、ユーザはあらゆる PC アプリケーションをフル パフォーマンスで実行することができます。

3.2 GRID vApps

GRID vApps は、リモート デスクトップ セッション ホスト (RDSH) またはその他のアプリケーション ストリーミングやセッションベースのソリューションを導入する組織向けの製品です。PC の Windows アプリケーションをフル パフォーマンスで実行できるように設計された Windows Server ホスト型の RDSH デスクトップも、GRID vApps でサポートされています。

3.3 Quadro vDWS

このエディションは、Dassault CATIA、SOLIDWORKS、3DExcite、Siemens NX、PTC Creo、Schlumberger Petrel、Autodesk Maya などの強力な 3D コンテンツ作成アプリケーションを使用するメインストリームおよびハイエンドのデザイナーに最適です。NVIDIA Quadro Virtual Data Center Workstation を使用すると、プロフェッショナルなグラフィックス アプリケーションのすべての機能にフル パフォーマンスであらゆるデバイスからアクセスできます。

3.4 NVIDIA vCS

多くの組織が、人工知能 (AI) 、ディープ ラーニング (DL) 、データ サイエンスなど、コンピューティング負荷の高いサーバ ワークロードを実行しています。このようなユースケースでは、NVIDIA vComputeServer ソフトウェアが NVIDIA GPU を仮想化することで、ECC、ページ リタイアメント、NVLink 経由のピアツーピア、マルチ vGPU などの機能を使用して、コンピューティング負荷の高いサーバ ワークロードを高速化します。

注：Quadro vDWS ライセンスで GRID vPC と NVIDIA vCS を使用できます。

4 VMware vSphere 上の 3D ワークロード

VMware vSphere は、基盤となるリソースを効果的に活用し、アプリケーションに高可用性を提供することで、仮想化プラットフォームとしての信頼を獲得しました。最新の GPU ドライバを NVIDIA サイトからダウンロードして vSphere ホストにインストールできます。

注：VMware Update Manager を使用すると、複数のホストに NVIDIA vGPU ソフトウェアを導入できます。オフライン バンドルの.zip ファイルを使用して、「Host Extension」タイプのベスラインを作成してください。

VM は次のいずれかで GPU リソースを消費します。

- Virtual Dedicated Graphics (vDGA)
- Virtual Shared Graphics (vSGA)
- Virtual Shared Passthrough Graphics (NVIDIA vGPU)

4.1 Virtual Dedicated Graphics

VM は、Direct IO を使用する PCI パススルー オプションを使用して、GPU にアクセスできます。ただし、スナップショットや vMotion など一部の vSphere 機能はサポートされません。VM ではネイティブ レベルのパフォーマンスが実現します。

1. vDGA を設定するには、[Hardware]セクションの[PCI Devices]をクリックし、[Configure Passthrough]をクリックします。

Summary Monitor **Configure** Permissions VMs Datastores Networks More Objects Updates

VMware File Location

System

- Licensing
- Host Profile
- Time Configuration
- Authentication Serv...
- Certificate
- Power Management
- Advanced System S...
- System Resource Re...
- Firewall
- Services
- Security Profile
- System Setup

PCI Devices

Postthrough-enabled devices All PCI devices

ID	Status	Vendor Name	Device Name
0000:18:00.0	Available	NVIDIA Corporation	NVIDIATesla T4
0000:AF:00.0	Available	NVIDIA Corporation	NVIDIATesla T4
0000:D8:00.0	Available	NVIDIA Corporation	NVIDIATesla T4

REFRESH CONFIGURE PASSTHROUGH

2. 次に、[Edit Settings]ページで、PCI デバイスを追加します。

Edit Settings MultiGPU

Virtual Hardware VM Options

ADD NEW DEVICE

> CPU	4	
> Memory	48	
> Hard disk 1	80	
> Hard disk 2	250	
> SCSI controller 0	VMware Paravirtual	
> Network adapter 1	VDS-01-VLAN3403	<input checked="" type="checkbox"/> Connect...
> CD/DVD drive 1	Client Device	<input type="checkbox"/> Connect...
> USB xHCI controller	USB 3.0	
> PCI device 0	0000:18:00.0 NVIDIATesla T4 NVIDIA Corporation	
> New PCI device	0000:af:00.0 NVIDIATesla T4 NVIDIA Corporation	
> New PCI device	0000:18:00.0 NVIDIATesla T4 NVIDIA Corporation	
> Video card	0000:18:00.0 NVIDIATesla T4 NVIDIA Corporation	
> VMCI device	0000:d8:00.0 NVIDIATesla T4 NVIDIA Corporation	

Device on the virtual machine PCI bus that provides support for the virtual machine communication interface

CANCEL OK

3. GPU が vDGA モード用に設定されている場合、グラフィック デバイスのリストの[Active Type]には[Direct]と表示されます。

vSphere 6.7 Update 2 までは、表示される 3 つのオプション (vDGA、vSGA、vGPU) のうち、VM が複数の GPU にアクセスできるのは vDGA だけでした。vSphere ホストでは NVIDIA vGPU ドライバはオプションですが、VM 内には OS ドライバが必要です。ライセンスをチェックイン / チェックアウトするために、NVIDIA ライセンス サーバが必要です。ライセンスがない場合は、使用できる機能が大幅に制限されます。

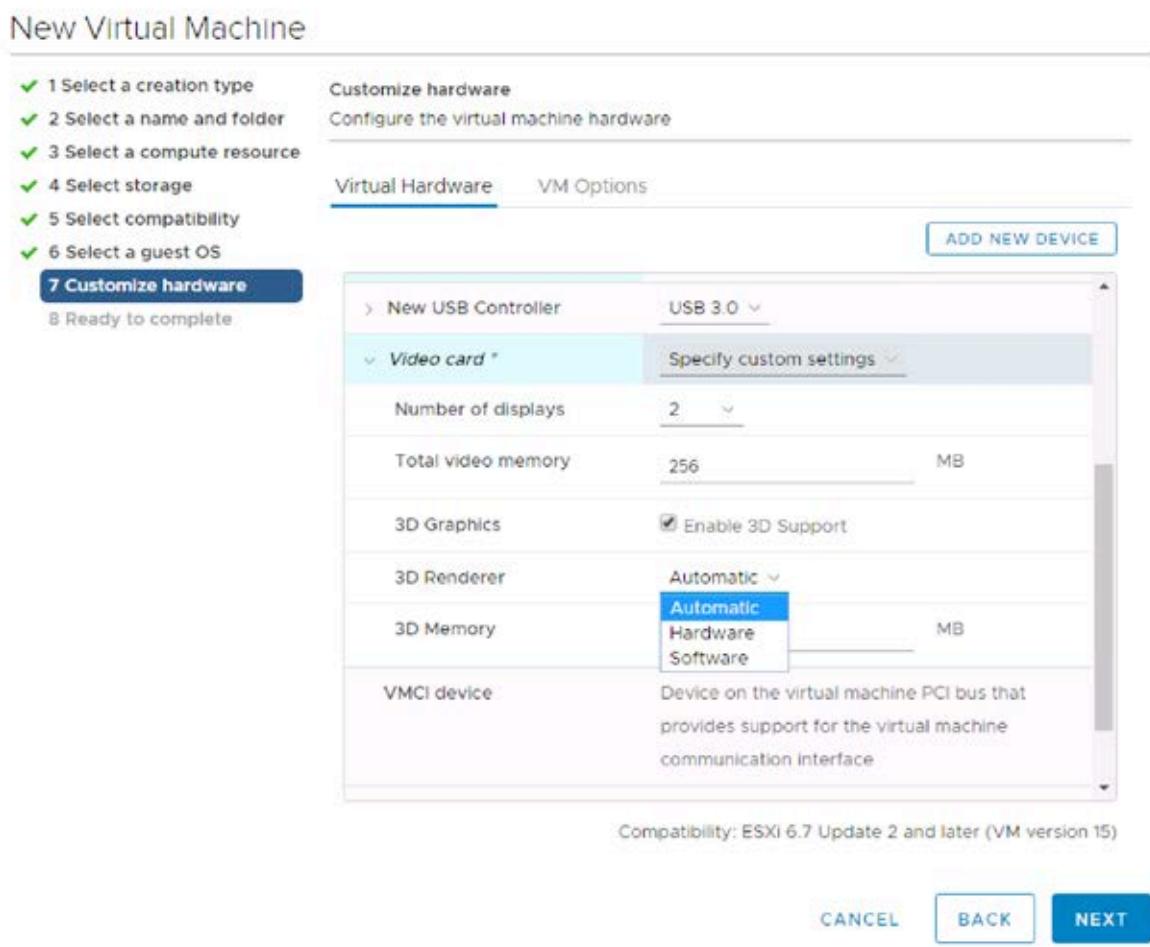
表 1) vDGA – H610C と H615C の比較

項目	H610C	H615C
サーバ 1 台あたりの最大 バススルーハイブ数	8	3

4.2 Virtual Shared Graphics

Virtual Shared Graphics は、VMware vSphere で有効になっているデフォルト モードです。ハードウェアベースのレンダリングを使用するためには、vSphere ホストに NVIDIA vGPU ソフトウェアが必要です。GPU がない場合、CPU サイクルを使用してソフトウェアベースのレンダリングが行われます。Virtual Shared Graphics を使用するには、3D サポートを有効にして、VMware Tools を VM にインストールします。

図 4) VM 上の 3D グラフィックス



VMware デバイス ドライバでは、DirectX と OpenGL のサポートに制限があります。また、4K モニタについても同様です。フレーム バッファ メモリは 2GB に制限されています。

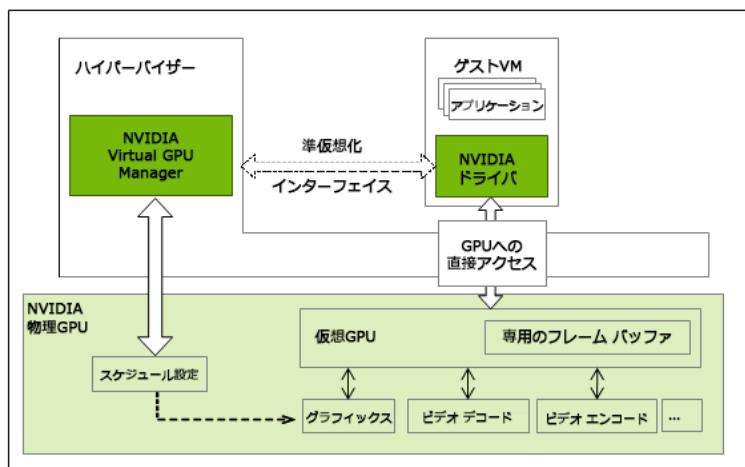
EVC が有効な場合、Virtual Shared Graphics モードを使用する vSphere クラスタに H610C ノードと H615C ノードを混在させることができます。

vDGA は、GPU カードから特定の VM への接続に PCI パススルーツを使用します。vDGA は vSphere の多くの機能をサポートしておらず、リソース利用率が低いため、ネットアップでは推奨していません。vSGA を使用すると、複数の VM が、ESXi ホストにローカルにインストールされた物理 GPU を利用してハードウェア アクセラレーションによる 3D グラフィックスを提供することができます。vSGA はグラフィックス API のサポートに制限があり、DirectX と OpenGL のすべてのバージョンがサポートされているわけではないため、ネットアップでは推奨していません。また、vSGA には CUDA のサポートもありません。vSGA を使用する VDI インスタンスは、ハイパー バイザーで実行されている Xorg サーバ経由でアクセスする VMware vSGA ドライバを使用しています。これは、パフォーマンスの観点からは最適な構成ではありません。仮想 GPU の場合は、全体を通して NVIDIA テクノロジが使用され、VDI インスタンスはネイティブの NVIDIA ドライバを実行する場合とほぼ同等のパフォーマンスを実現します。

Virtual Shared Passthrough Graphics

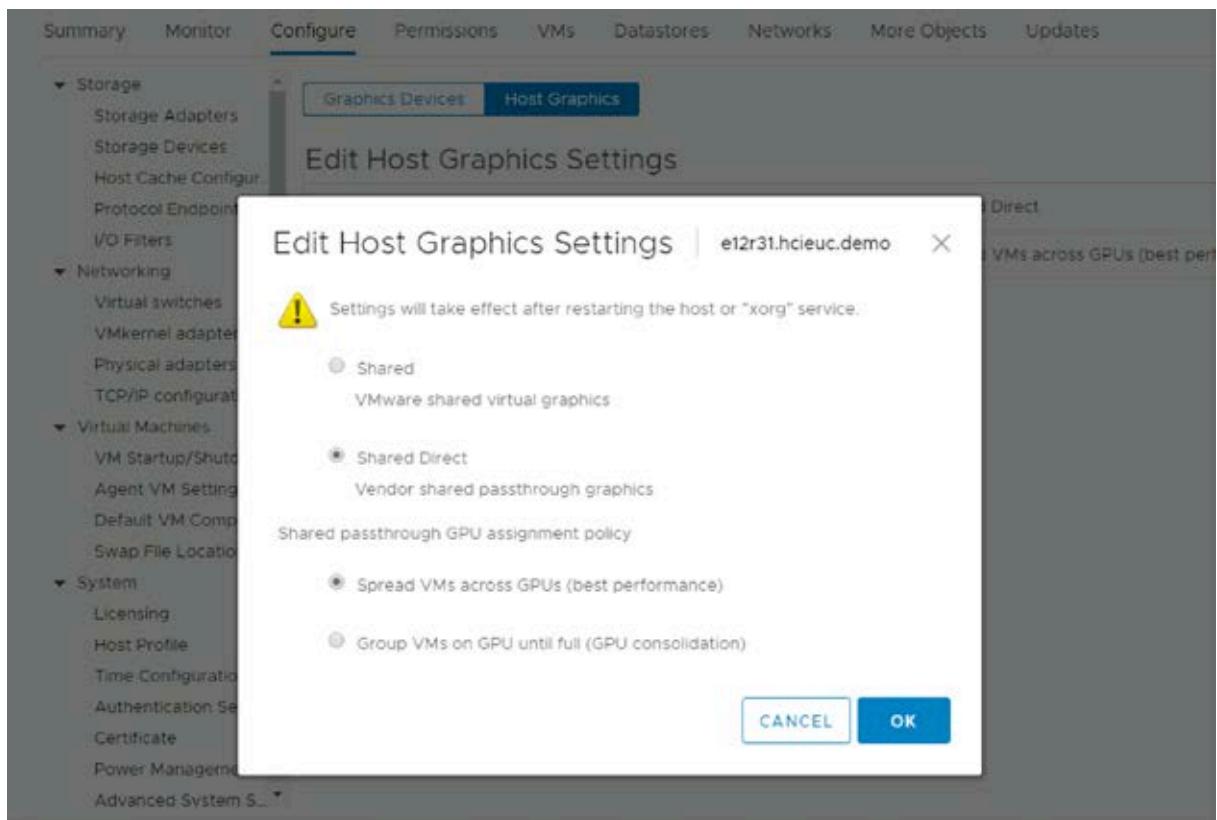
Virtual Shared Passthrough Graphics を使用すると、GPU リソースの利用率が向上します。各 VM には専用のフレーム バッファがあります。ただし、GPU のコンピューティング、エンコーダ、デコーダなどは共有されます。ハイパー バイザーが VM とリソースを共有するように、NVIDIA GRID ソフトウェアは VM からの GPU 要求に対してリソースのスケジュールを管理します。

図 5) NVIDIA vGPU のアーキテクチャ



Virtual Shared Passthrough Graphics モードを使用するには、ホスト グラフィックス設定を [Shared Direct] モードに変更する必要があります (図 6) を参照)。

図 6) [Shared Direct] モード



[VM Settings] で、共有 PCI デバイスを追加し、フレーム バッファに基づいて必要な vGPU プロファイル (表 2) を選択します。PCI デバイスに対する直接アクセスを可能にするためには、メモリを確保する必要があります。表 2 は、このあとのスクリーンショットで選択可能な T4 vGPU プロファイルの詳細です。

表 2) T4 vGPU プロファイル

仮想 GPU タイプ	ユースケース	フレーム バッファ (MB)	仮想 ディスプレイ ヘッド	ディスプレイ ヘッドあたりの最大解像度	GPUあたりの最大 vGPU 数	H615C サーバあたりの最大 vGPU 数	必要なライセンス エディション
T4-16Q	仮想ワークステーション	16384	4	4096x2160	1	3	Quadro vDWS
T4-8Q	仮想ワークステーション	8192	4	4096x2160	2	6	Quadro vDWS
T4-4Q	仮想ワークステーション	4096	4	4096x2160	4	12	Quadro vDWS
T4-2Q	仮想ワークステーション	2048	4	4096x2160	8	24	Quadro vDWS
T4-1Q	仮想デスクトップ、仮想ワークステーション	1024	2	4096x2160	16	48	Quadro vDWS
T4-16C	トレーニング ワークロード	16384	1	4096x2160	1	3	vComputeServer または Quadro vDWS
T4-8C	トレーニング ワークロード	8192	1	4096x2160	2	6	vComputeServer または Quadro vDWS
T4-4C	推論ワーク ロード	4096	1	4096x2160	4	12	vComputeServer または Quadro vDWS
T4-2B	仮想デスクトップ	2048	2	4096x2160	8	24	GRID Virtual PC または Quadro vDWS
T4-2B4	仮想デスクトップ	2048	4	2560x1600	8	24	GRID Virtual PC または Quadro vDWS
T4-1B	仮想デスクトップ	1024	4	2560x1600	16	48	GRID Virtual PC または Quadro vDWS

仮想 GPU タイプ	ユースケース	フレーム バッファ (MB)	仮想 ディスプレイ ヘッド	ディスプレイ ヘッドあたりの最大 解像度	GPUあたりの最大 vGPU 数	H615C サーバあたりの最大 vGPU 数	必要なライセンス エディション
T4-1B4	仮想デスクトップ	1024	1	4096x2160	16	48	GRID Virtual PC または Quadro vDWS
T4-16A	仮想アプリケーション	16384	1	1280x1024	1	3	GRID Virtual Application
T4-8A	仮想アプリケーション	8192	1	1280x1024	2	6	GRID Virtual Application
T4-4A	仮想アプリケーション	4096	1	1280x1024	4	12	GRID Virtual Application
T4-2A	仮想アプリケーション	2048	1	1280x1024	8	24	GRID Virtual Application
TT4-1A	仮想アプリケーション	1024	1	1280x1024	16	48	GRID Virtual Application

NVIDIA vGPU モードの場合は、VM で 3D サポートを有効にしないでください。詳細は、NVIDIA vGPU のユーザ ガイドを参照してください。

GRID Virtual PC (プロファイルの末尾は通常「B」) は仮想デスクトップに、GRID Virtual Application (末尾が「A」のプロファイル) はホスト型アプリケーションに使用されます。ほとんどの「Q」プロファイル (Quadro vDWS) は最大 4 台の 4K モニタをサポートします。これにより、画像編集ツールのユーザ エクスペリエンスが向上し、プロフェッショナルなグラフィックスワーク ロードへの対応が実現します。

NVIDIA T4 GPU は複数のプロファイルをサポートします。NVIDIA では、標準的なナレッジ ワー カーに最適なエクスペリエンスを提供するためには、GRID Virtual PC (GRID vPC) の 1GB プロファイルを推奨しています。

- ブラウジング、E メールの送受信、複雑なドキュメント、プレゼンテーション、スプレッドシートの作成など、使用頻度が高いアプリケーション
- Microsoft Windows 10、Microsoft Office 生産性アプリケーション、ストリーミングビデオ、最新の Web 標準 (WebGL など) を使用したマルチメディアなどのアプリケーション
- 最大 4 台の 2K (2560x1600) モニタ

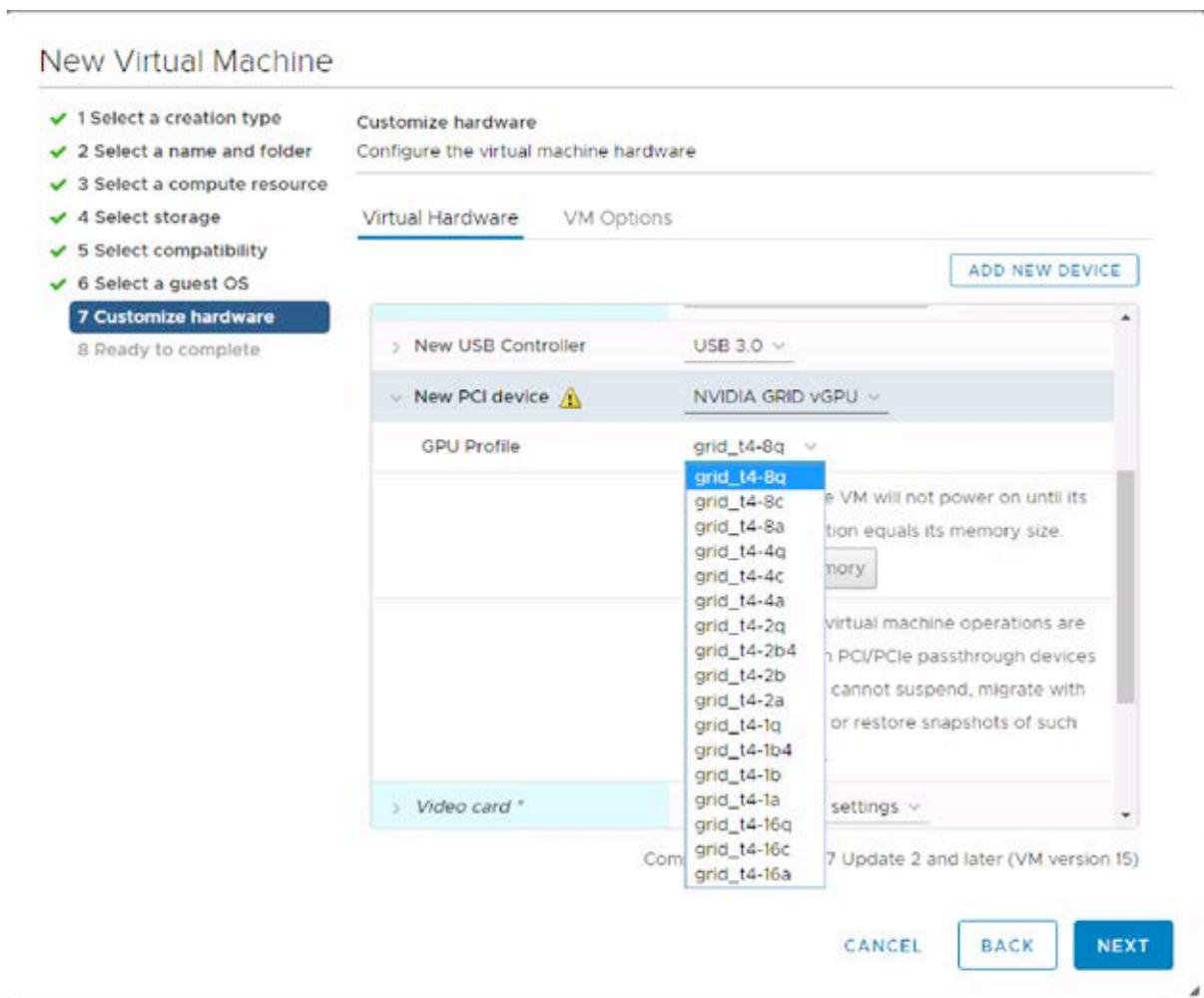
一方、以下のいずれかに該当するユーザには、高度なナレッジ ワーカー向けの GRID vPC 2GB プロファイルを割り当てる必要があります。

- 大容量のファイルや高解像度のメディアを処理するために複数の高解像度モニタが必要など、特定の要件があるユーザ
- Microsoft Windows 10、Microsoft Office 生産性アプリケーション、ビデオ、マルチメディア、業界固有アプリケーション（例：Bloomberg、Thomson Reuters Eikon、DICOM のビューア）などのアプリケーション
- 最大 2 台の 4K (4096x2160) モニタ

Dassault Systèmes CATIA、Autodesk Revit、Siemens NX、Petrel などのアプリケーションを実行するプロのクリエイターやテクニカル プロフェッショナルは、操作するモデルのワークフローとサイズによっては多くのフレーム バッファを必要とします。

AI、機械学習、データ サイエンスなどのコンピューティング ワークフローに対して、NVIDIA は 4GB 以上のフレーム バッファを推奨しています。モデルの規模が大きいほど、推奨されるフレーム バッファ サイズも大きくなります。

図 7) GPU プロファイルの選択



NVIDIA は、GPU で使用可能なものと同じ vGPU プロファイルをサポートします（表 3）。

表 3) 単一の GPU における NVIDIA T4 vGPU プロファイル

Tesla T4															
T4-16Q															
T4-8Q								T4-8Q							
T4-4Q				T4-4Q				T4-4Q				T4-4Q			
T4-2Q		T4-2Q		T4-2Q		T4-2Q		T4-2Q		T4-2Q		T4-2Q		T4-2Q	
T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q	T4-1Q

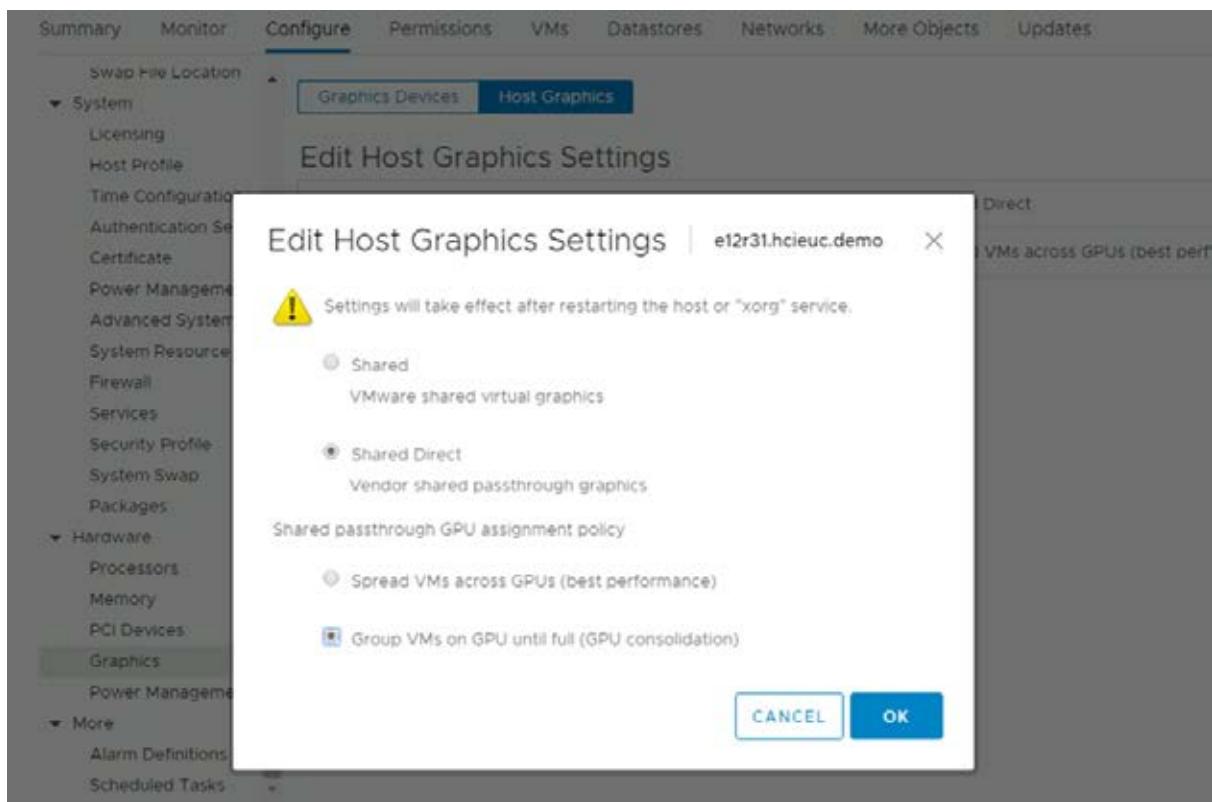
NVIDIA vGPU ソフトウェアでは、GPU を 1 つだけ使用する GPU アクセラレータに複数のプロファイルを使用できません。たとえば、4Q プロファイルを使用する VM が GPU で実行されている場合、その GPU でホストできるのは同じ 4Q プロファイルを使用する別の VM だけです。

表 4) 単一の GPU における vGPU プロファイルの混在

T4-4Q	T4-2Q	T4-2Q
-------	-------	-------

H610C 上の vGPU プロファイルは H615C 上のプロファイルとは異なるため、VM をホスト間で移行することはできません。そのため、クラスタ内では同じモデルを使用することを推奨します。クラスタ内でどうしてもノードを混在させる必要がある場合は、GPU 割り当てポリシーを「Group VMs on GPU」に変更することを検討してください（図 8）。どちらかの GPU ノードで障害が発生した場合のテイクオーバー用に十分なリソースが必要です。

図 8) [Group VMs on GPU until full]



NVIDIA ドライバを VM にインストールすると、VM コンソールには vGPU プロファイル用の空白の画面が表示されます。コンソールにアクセスするには、NVIDIA ドライバの前に VNC / Horizon Direct Connect をインストールする必要があります。

VMware vSphere 6.7 Update 1 以降では、NVIDIA vGPU プロファイルを使用する VM で vMotion がサポートされます。

表 5) vGPU – H610C と H615C の比較

フレーム バッファ	H610Cx1 (M10x2) 2RU	H615Cx1 (T4x3) 1RU	H615Cx2 (T4x6) 2RU
1GB	64	48	96
2GB	32	24	48
4GB	16	12	24
8GB	8	6	12
16GB	NA	3	6

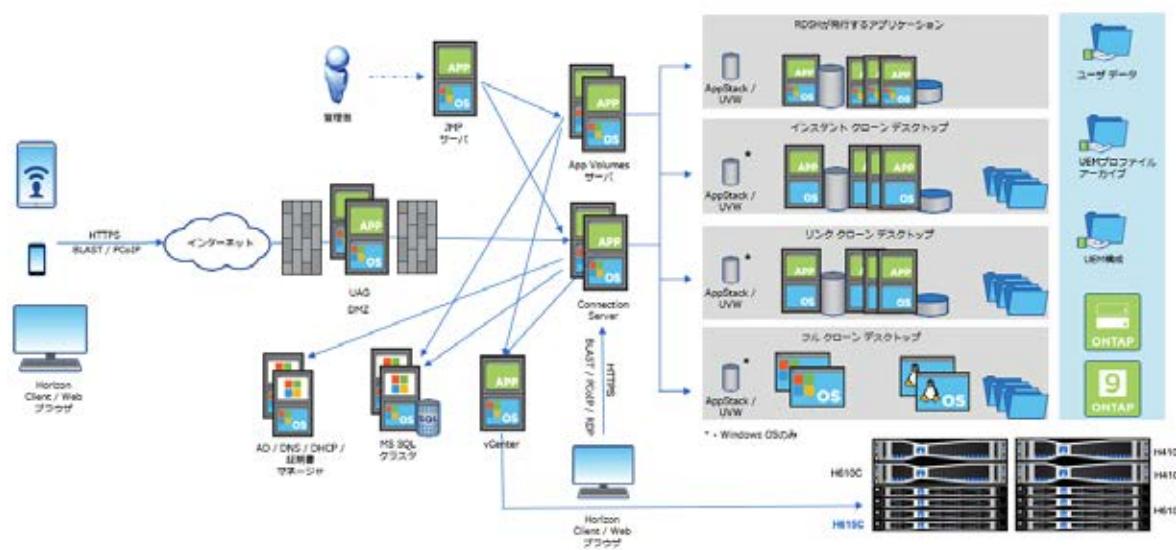
H615C は、同じサイズのラック ユニットで 2 倍の台数をホスト可能なため、ほとんどのワークloadsに対してパフォーマンスも倍増します。

5 VMware Horizon

5.1 ソリューション リファレンス アーキテクチャ

図 9 は、NetApp HCI を使用した VMware Horizon の代表的なリファレンス アーキテクチャを示しています。

図 9) ソリューション リファレンス アーキテクチャ



5.2 VMware Horizon Connection Server

VMware Horizon は、仮想デスクトップとホスト型アプリケーションのコネクション ブローカーです。VMware Horizon では、デスクトップのフル クローン、リンク クローン、またはインスタント クローンの自動プロビジョニングを実行できます。また、仮想マシンだけでなく物理マシンも管理できます。サーバベースのコンピューティングは、リンク クローンまたはインスタント クローンを使用して自動でプロビジョニング可能な RDSH によって実現します。

ナレッジ ワーカー、または病院やキオスクなどのシナリオでは、アプリケーションへの瞬時のアクセスが必要になることがあります。このようなシナリオでは、あらゆるデバイスでネイティブ PC 並みのエクスペリエンスを提供することが重要です。この種のユースケースでは、リモート デスクトップ セッション サーバ フーム上のホスト型アプリケーションが役に立ちます。VMware Horizon 7.9 から、VM でホストされるアプリケーションがサポートされるようになり、Windows 10 のユニバーサル Windows プラットフォーム (UWP) アプリケーションをホストできるようになりました。

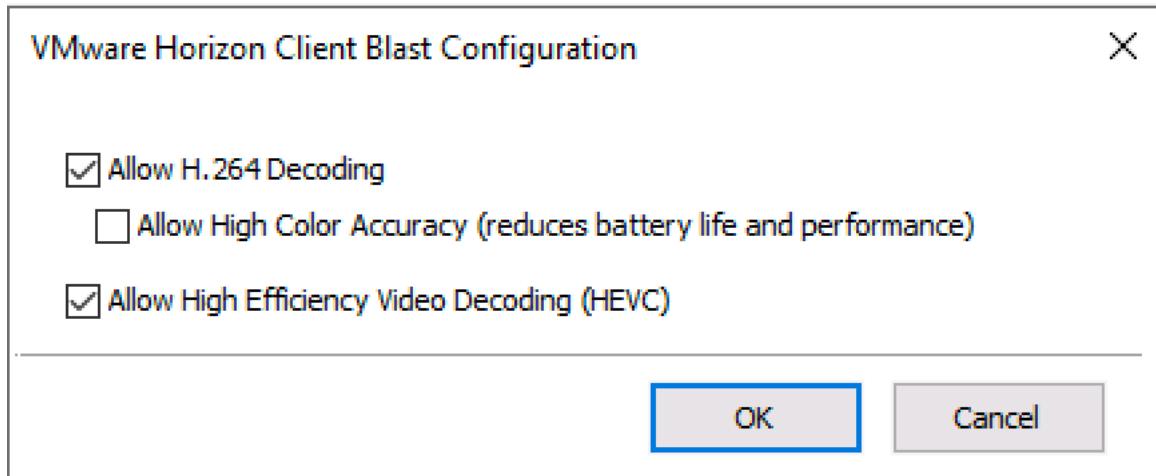
注 : VMware Horizon 7.9 には、NVIDIA T4 のサポートが追加されました。7.9 は、NetApp HCI H615C と一緒に使用する場合の必要最小バージョンです。

ホスト型アプリケーションを使用すると、ハイ レンダリングを必要とするアプリケーションを数台の H615C ホストに統合できます。また、Windows 上のアプリケーション ホストにアクセスする必要のある Linux ユーザにもホスト型アプリケーションが役に立ちます。

5.3 Horizon Client

VMware Horizon は、RDP、PCoIP、BLAST などのさまざまな表示プロトコルをサポートします。H.264 での BLAST Extreme のサポートはデフォルトで有効になっています。VMware Horizon は、ハードウェア デコーダがある場合はハードウェア デコーダを使用、ない場合はソフトウェアベースのデコードを使用します。H.265 (別名 HEVC) は Horizon Client ではオプションです。HEVC を使用すると、高解像度ディスプレイのネットワーク トラフィックが削減されます。HEVC を有効にする方法については、図 10 を参照してください。

図 10) HEVC の有効化



HTML クライアントでは、HEVC をサポートする Google Chrome ブラウザを使用してください。仮想デスクトップにインストールされた Horizon Agent はエンコードを実行し、エンド デバイス（スマートフォン、タブレット、ラップトップなど）上の Horizon Client はデコードを実行します。

5.4 App Volumes

管理する VM テンプレートとイメージの数は、さまざまなオペレーティング システム イメージやパッチ、複数のバージョンのアプリケーションの存在によって短期間で増加する傾向にあります。App Volumes はアプリケーションとユーザ データをオペレーティング システム イメージから切り離し、アプリケーションは仮想ディスクに格納されます。これらの仮想ディスクは、vSphere ホストに提示されるボリュームからマウントすることも、VM 内に直接マウントすることもできます。App Volumes の設定は Microsoft SQL データベースにホストされ、複数の App Volumes サーバがデータにアクセスして高可用性を提供できます。

5.5 User Environment Manager

User Environment Manager (UEM) は、ユーザ プロファイルとユーザ アプリケーション設定（アプリケーションのレイアウト、外観、言語設定など）を管理します。UEM の設定は SMB ファイル共有に格納されます。VMware Horizon のクライアント ポリシーはさまざまな条件を使用して適用でき、条件は UEM 内で定義することができます。たとえば、オフィスでは USB リダイレクトを許可し、オフィスの外では許可しないように設定できます。

5.6 SMB ファイル共有

UEM は、SMB ファイル共有を使用してユーザ プロファイルをバックアップし、また構成ファイルをホストします。App Volumes は、仮想ディスクを VM 内にマウントする必要がある場合に SMB 共有を必要とします。NetApp HCI は NetApp ONTAP® Select を使用してファイル サービスを提供します。ONTAP Select は VMware Horizon の実装に役立つ次の機能を提供します。

- ユーザのホーム ディレクトリ
- SVM (Storage Virtual Machine)
- NetApp FlexGroup テクノロジ
- NetApp FabricPool テクノロジ
- NetApp Data Availability Services
- アダプティブ QoS
- 重複排除
- データ保護機能
 - NetApp RAID DP® テクノロジ

- NetApp SnapshotTM コピー
- NetApp FlexClone[®] テクノロジ
- NetApp SnapMirror[®] / SnapVault[®] テクノロジ
- Snapshot コピーからのセルフ リストア

5.7 Just-in-Time Management Platform

Just-in-Time Management Platform (JMP) サーバは、VMware Horizon の新しい管理インターフェイスをホストします。このインターフェイスを使用して、デスクトップとアプリケーションのプロビジョニング、ユーザとグループの割り当て、および UEM で定義されたポリシーの関連付けを行うことができます。

5.8 Universal Access Gateway

Universal Access Gateway を使用すると、ユーザは外部ネットワークから仮想デスクトップやアプリケーションに安全にアクセスできます。

5.9 VMware vRealize Log Insight

VMware vRealize Log Insight では、ログ ファイルを一元的に収集してインデックスを作成します。管理者は、複数のログ ファイルをすばやく検索して、潜在的な問題を特定できます。

5.10 VMware vRealize Operations

VMware vRealize Operations (vROps) は、インフラ コンポーネントを一元的に監視します。vROps 用の NetApp HCI プラグインは Blue Medora から入手でき、NVIDIA は NVIDIA GPU を監視するための管理パックを提供しています。vROps 用の NVIDIA 仮想 GPU による管理サポートにより、クラスタ、ホスト、または VM レベルで GPU 対応環境の全体像を把握でき、リアルタイムでの監視とプロアクティブな管理が実現します。

6 SPECviewperf 13 ベンチマーク

6.1 概要

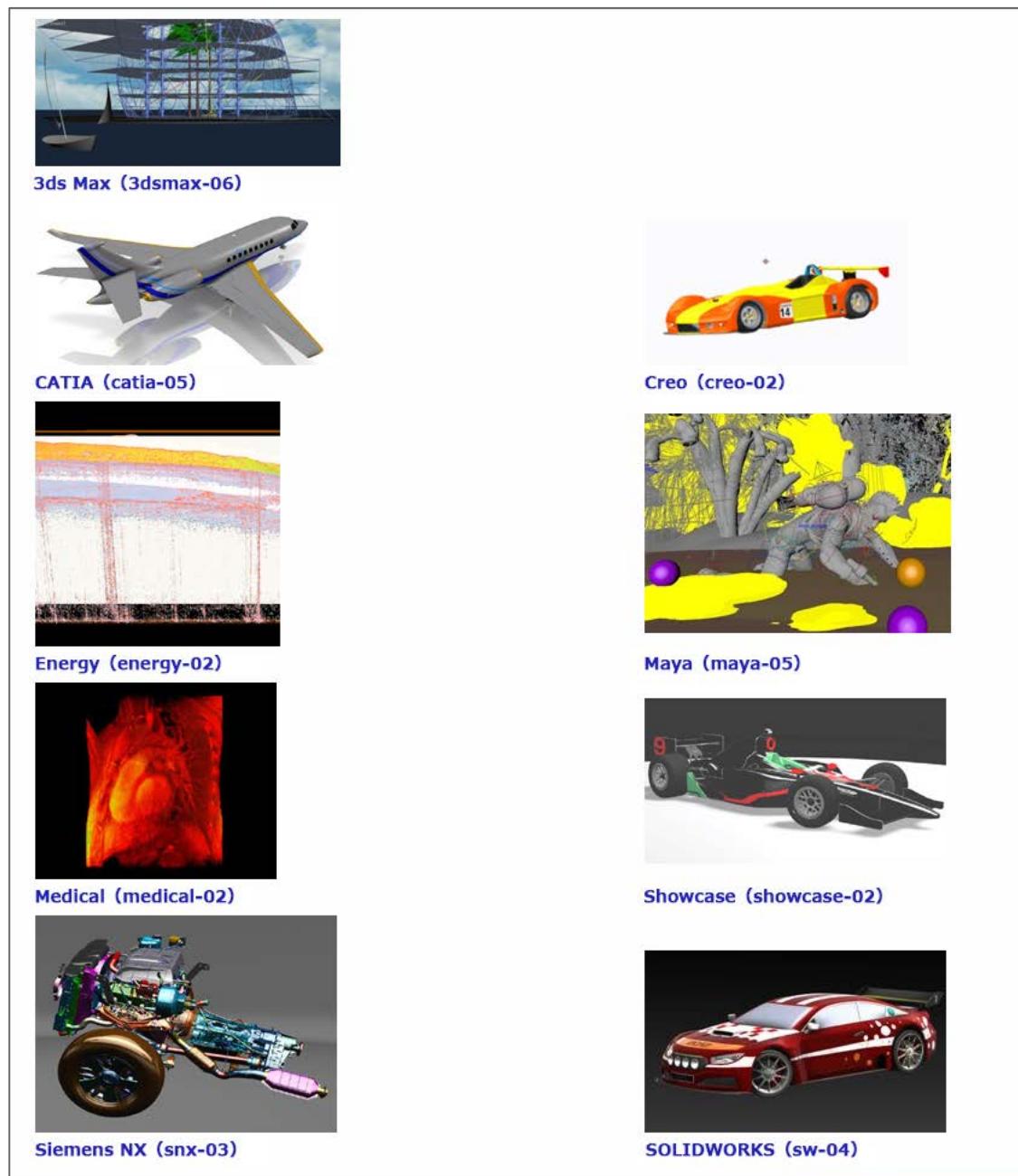
Standard Performance Evaluation Corporation (SPEC) は、ベンチマークとツールの標準化と維持、および推奨を目的に設立された非営利団体です。グラフィックスとワークステーションのパフォーマンスについては Application Performance Characterization (SPECapc) ツールを推奨しています。SPECapc は特定のツールを対象とし、ベンダー ライセンスが必要です。SPECapc ツールには次のサンプル リストがあります。

- SPECapcSM for 3ds Max 2015
- SPECapcSM for Maya 2017
- SPECapcSM for PTC Creo 3.0
- SPECapcSM for Siemens NX 9.0 / 10.0
- SPECapcSM for SOLIDWORKS 2017

SPECviewperf と SPECworkstation はどちらも、専門的なアプリケーションに基づいてグラフィックス パフォーマンスを測定します。これらのベンチマークでは、OpenGL と Direct X アプリケーション プログラミング インターフェイスを使用して 3D グラフィックス パフォーマンスを測定します。ベンチマークのワークロードはビューセットと呼ばれ、実際のアプリケーションのグラフィックス コンテンツと動作をシミュレートします。

SPECviewperf は GPU カードを対象としています。SPECworkstation は、ワークステーションの主要な要素 (CPU、メモリ、ストレージなど) もすべて測定します。SPECworkstation のテスト期間は SPECviewperf よりも長くなります。図 11 は、SPECviewperf 13 のビューセットを示しています。

図 11) SPECviewperf 13 のビューセット



注：3ds Max ビューセットと Showcase ビューセットは、4K テストには使用できません。

テストは NVIDIA nVector ツールセットを使用して実施しました。このツールセットは、VM とクラウドアント マシンの作成、エンコード オプションの設定、テストの実施、パフォーマンス データのキャプチャをオーケストレーションします。

テストに使用したハードウェアとソフトウェアについては、表 6 と表 7 を参照してください。

表 6) ハードウェア リスト

モデル	数	説明
H410S	4	ストレージ クラスタ
H410C	2	インフラ クラスタ

モデル	数	説明
H615C (2x18C、382GB)	2	デスクトップ リソース プール
Mellanox SN2010 スイッチ	4	25Gb スイッチ

表 7) ソフトウェア リスト

製品	説明
Microsoft Windows Server 2016	AD、DNS、証明書サーバ等用のサーバ OS
Microsoft Windows Server 2019	VMware Horizon 7.9 用のサーバ OS
VMware vSphere 6.7 Update 2、ビルド 13006603	ハイパーバイザー
VMware vCenter 6.7 Update 2b	管理
NVIDIA vGPU Manager Windows ドライバ	9.0 430.27 431.02
VMware Horizon 7.9 VMware Horizon Client 5.1.0、ビルド 14045148	コネクション ブローカー デスクトップにアクセスするためのクライアント ソフトウェア
SPECviewperf 13	ベンチマーク ツール
Microsoft Windows 10 Pro Workstation バージョン 1903	デスクトップ OS

注：25Gb スイッチも使用可能でしたが、より一般的な 10Gb でテストを実施しました。

それぞれのビューセットについて、16Q プロファイルの VM 1 台（単一の T4 でフル フレーム バッファ）および 4Q プロファイルの VM 12 台（T4 カード 3 基を装着した H615C でフル ロード）を使用してテストを実施しました。ワークロード シミュレーション時の vSphere の CPU 利用率と GPU 利用率のパフォーマンス特性を以下に記載します。VM 1 台のテストでは、vCPU 18 基、16GB RAM の VM 構成を使用しました。VM 12 台のテストでは、vCPU 6 基、16GB RAM の VM 構成を使用しました。フレーム レート制限は無効にしました。クライアント VM の構成は、vCPU 4 基、4GB RAM、1Q プロファイルの GPU です。

SPECviewperf は、さまざまなアプリケーションと用途モデルを対象に、グラフィックス カードがシーンをレンダリングできるフレーム レート (FPS) を測定します。各ビューセットは個別のアプリケーションまたは用途モデルをシミュレートし、グラフ内の各総合スコアはさまざまなシーンやレンダリング モードの加重幾何平均に基づいています。

6.2 3ds Max (3dsmax-06)

3ds Max ビューセットは、AutoDesk 3ds Max 2016 で生成されたグラフィックス ワークロードのトレースから作成されました。ビューセット内のレンダリング スタイルは、主要な市場で最も一般的に使用されているスタイル（リアリスティック、シェーディング、ワイヤフレームなど）を反映しています。また、それほど一般的ではありませんが、特徴的なレンダリング モデル（フラット、グラファイト、クレイなど）も組み込まれています。ビューセット内のアニメーションは、モデルに応じて、モデルの回転とカメラのフライスルーを組み合わせたものです。

次のビューセット テストを実行しました。

- 建築モデル、シェーディング
- 建築モデル、グラファイト

- 空間モデル、ワイヤフレーム
- 空間モデル、クレイ
- 水中モデル、ワイヤフレーム
- 水中モデル、シェーディング
- ヒューフィッシュモデル、シェーディング
- オフィスモデル、リアリスティック
- オフィスモデル、シェーディング
- オフィスモデル、リアリスティック、マテリアルあり

詳細については、[3ds Max viewset \(3dsmax-06\) のページ](#)を参照してください。

図 12 に、SPECviewperf 13 3ds Max ビューセットの総合スコアを次に示します。

図 12) 3ds Max の総合スコア

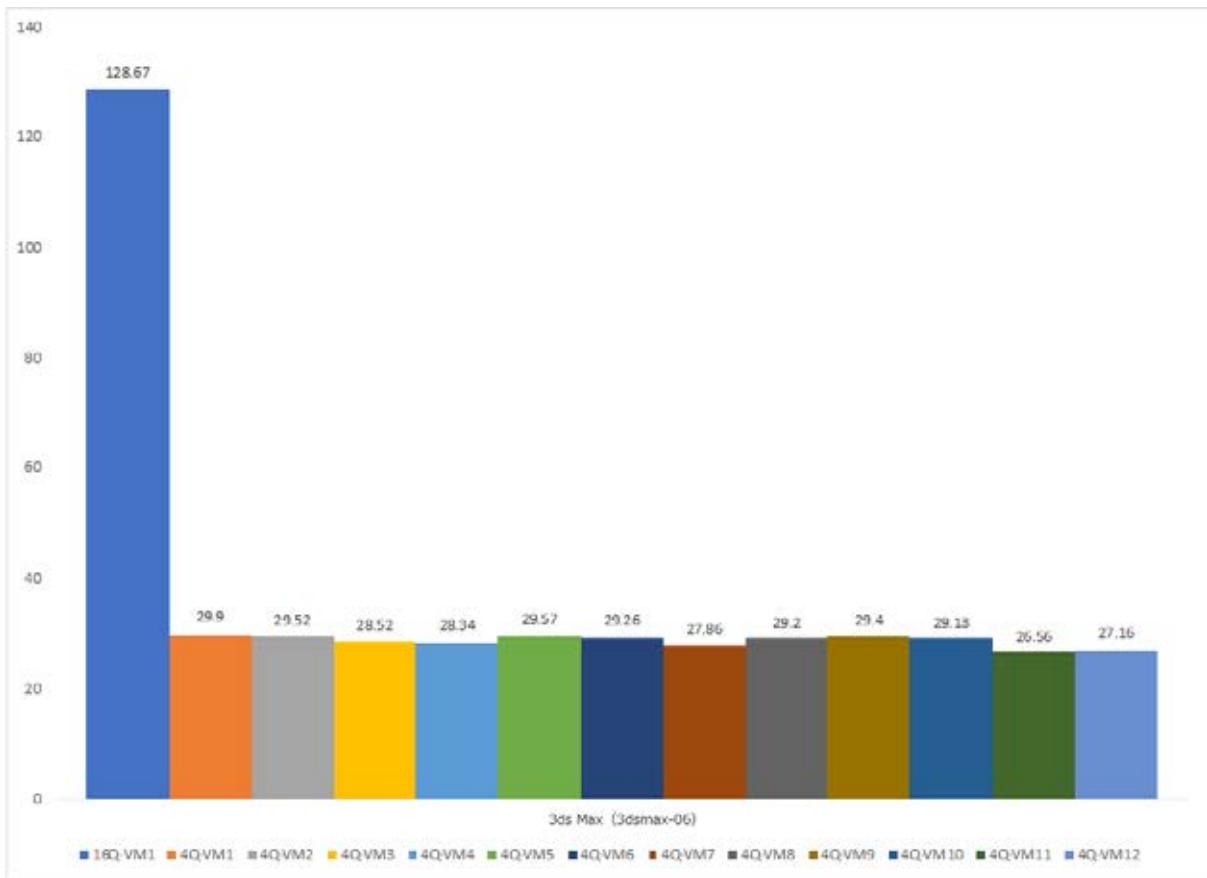
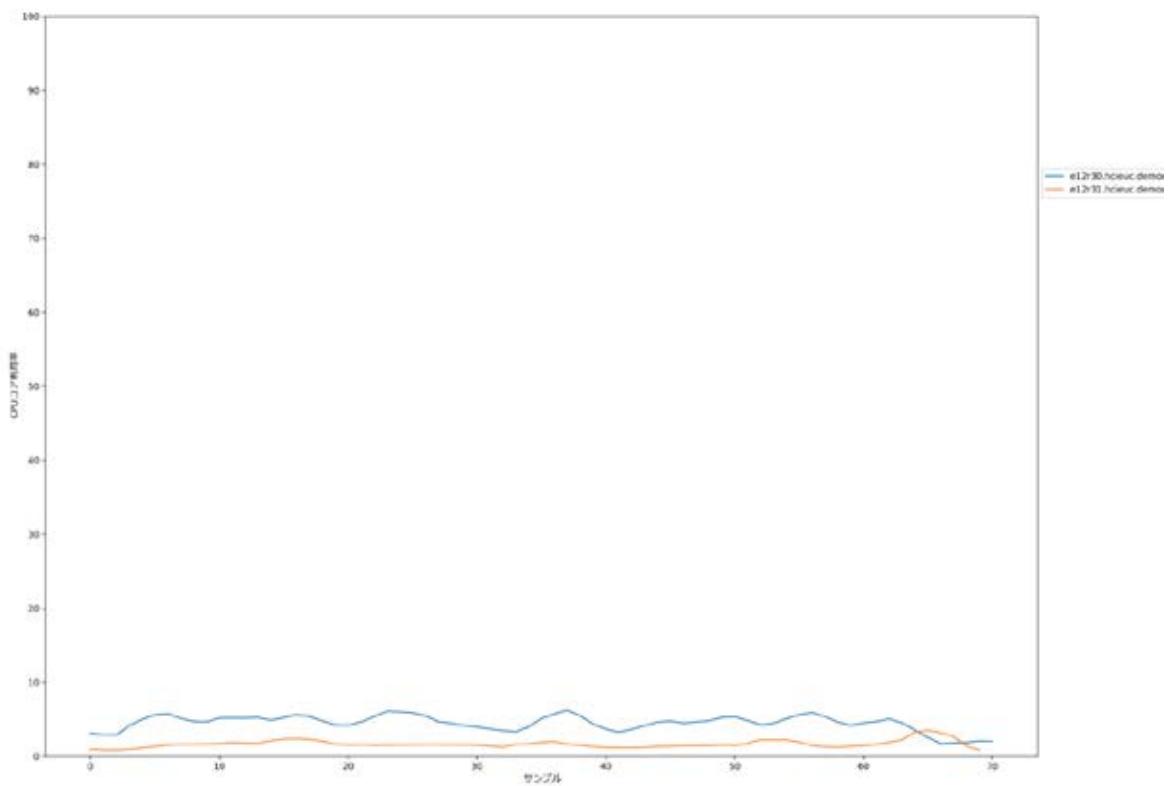


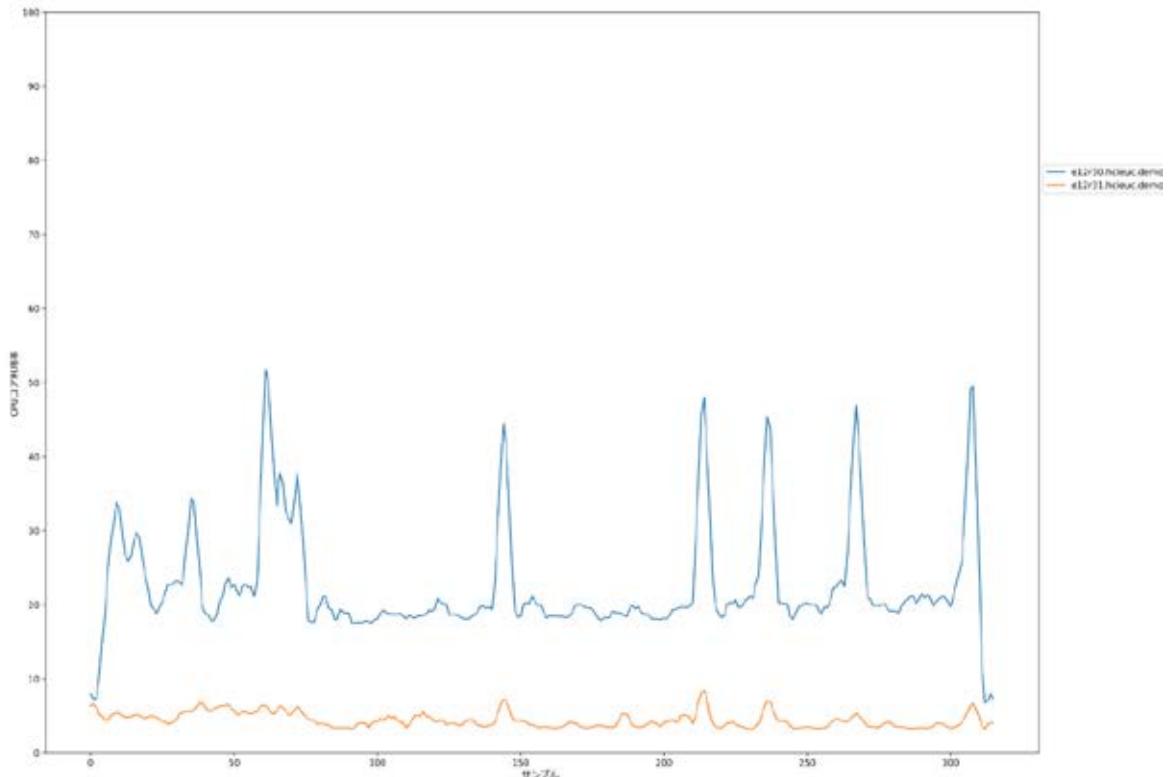
図 13 からわかるように、VM 1 台の vSphere ホストのコア利用率は 10% 以下です。ビューセットは e12r30 ホスト、クライアントは e12r31 ホストで実行しました。

図 13) 3ds Max : vSphere のコア利用率 – 16Qx1



VM 12 台の場合、vSphere のコア利用率は 60% 以下でした（図 14）。

図 14) 3ds Max : vSphere のコア利用率 – 4Qx12



VM 1 台および 12 台の場合の GPU 利用率を図 15 に示します。

図 15) 3ds Max : GPU 利用率 – 16Qx1

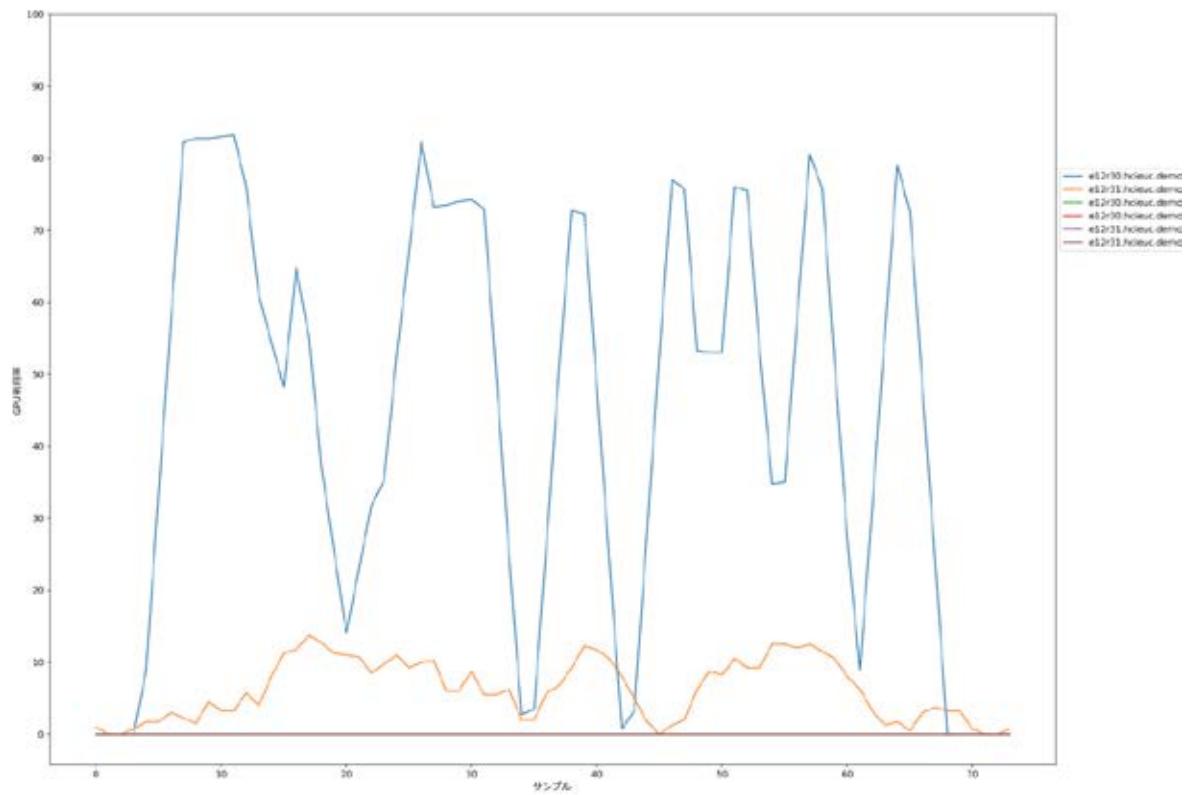
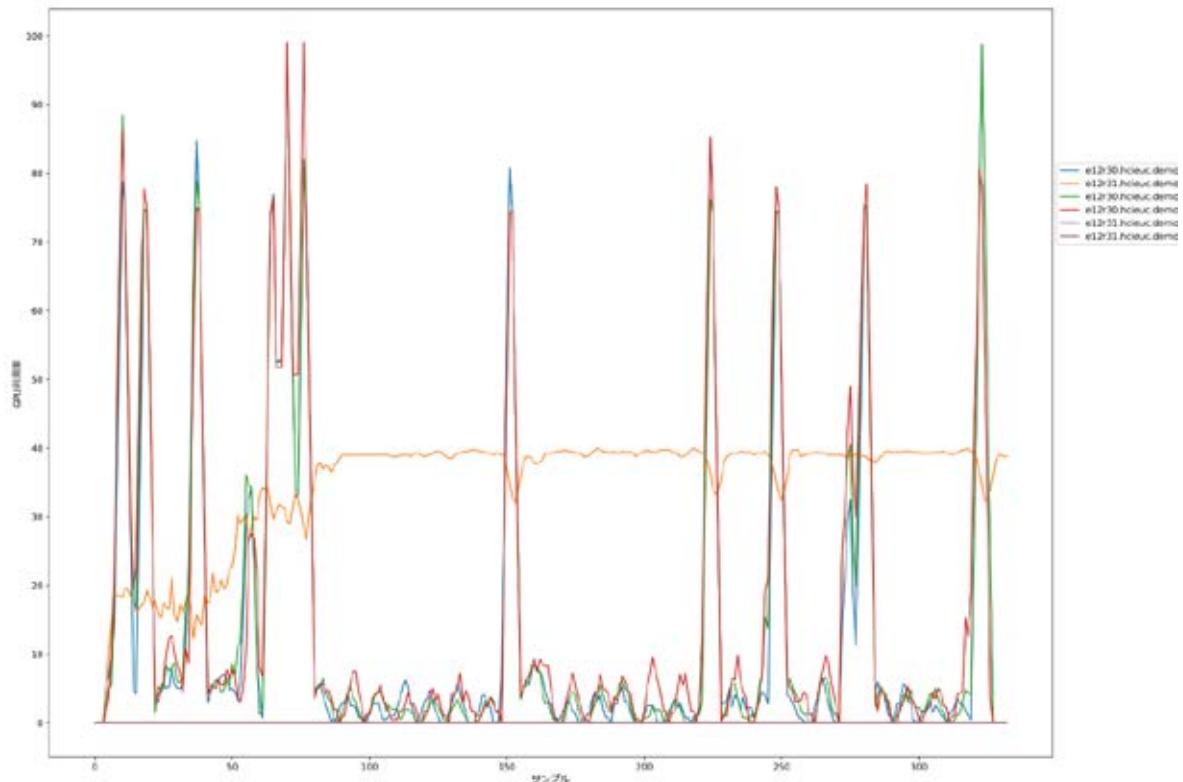


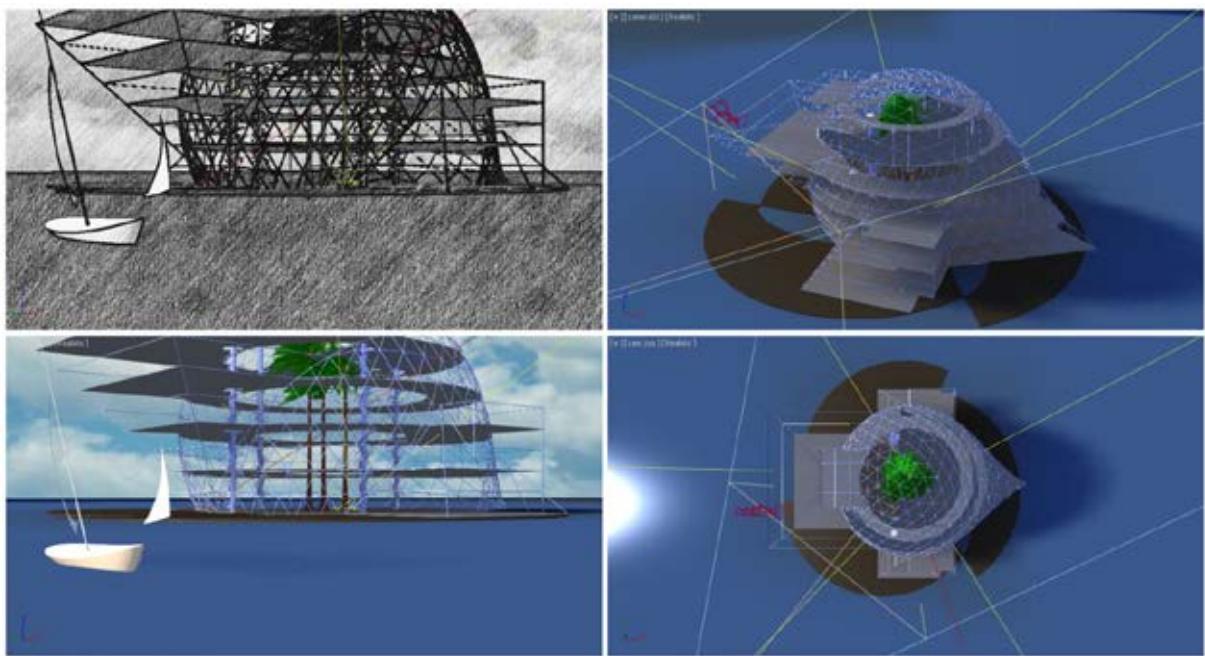
図 16) 3ds Max : GPU 利用率 – 4Qx12



深度優先の vGPU 割り当てポリシーでは、12 台のクライアント VM がすべて同じ GPU でホストされます。

図 17 は、テスト中にキャプチャしたサンプルスクリーンショットです。

図 17) 3ds Max のサンプル



6.3 CATIA (catia-05)

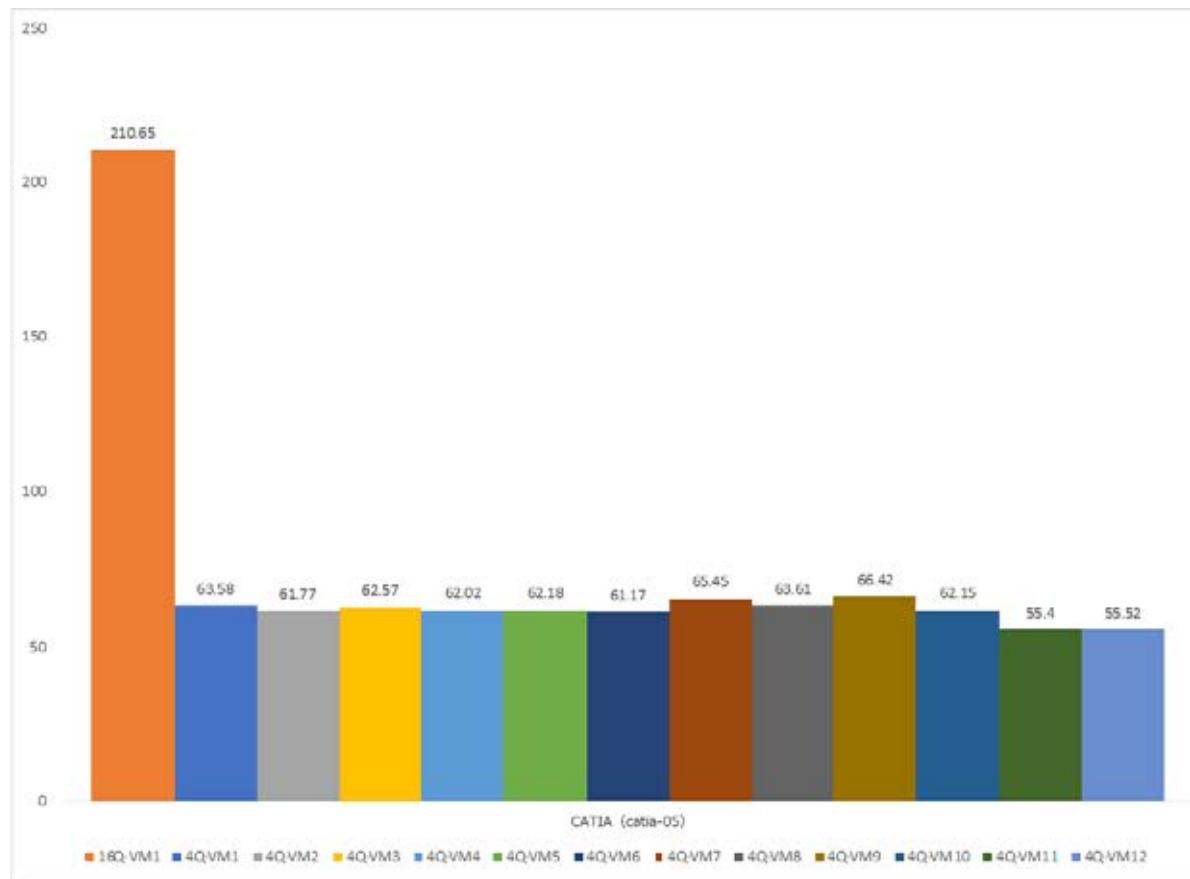
catia-05 ビューセットは、Dassault Systemes の CATIA V6 R2012 アプリケーションで生成されたグラフィックス ワークロードのトレースから作成されました。モデル サイズは頂点数 510 万～2,100 万です。

このビューセットには、ワイヤフレーム、アンチエイリアス、シェーディング、エッジ付きシェーディング、被写界深度、アンビエント オクルージョンなど、アプリケーションでサポートされる多数のレンダリング モードが含まれています。

次のビューセット テストを実行しました。

- レース カー (シェーディング、アンビエント オクルージョン、被写界深度エフェクトを適用)
- レース カー (シェーディング、鉛筆エフェクトを適用)
- レース カー (シェーディング、アンビエント オクルージョンを適用)
- 飛行機 (シェーディング、アンビエント オクルージョン、被写界深度エフェクトを適用)
- 飛行機 (シェーディング、鉛筆エフェクトを適用)
- 飛行機 (シェーディング)
- 飛行機 (エッジ付きのシェーディング)
- 飛行機 (シェーディング、アンビエント オクルージョンを適用)
- SUV1 車 (シェーディング、地表反射、アンビエント オクルージョンを適用)
- SUV2 車 (シェーディング、地表陰影を適用)
- SUV2 車 (シェーディング、地表反射、アンビエント オクルージョンを適用)
- ジェット機 (シェーディング、地表反射、アンビエント オクルージョンを適用)
- ジェット機 (エッジ付きのシェーディング、地表反射、アンビエント オクルージョンを適用)

図 18) CATIA の総合スコア



vSphere ホストのコア利用率は、VM 1 台で 10%以下 (図 19) 、VM 12 台で 80%以下 (図 20) でした。

図 19) CATIA : vSphere の CPU 利用率 – 16Qx1

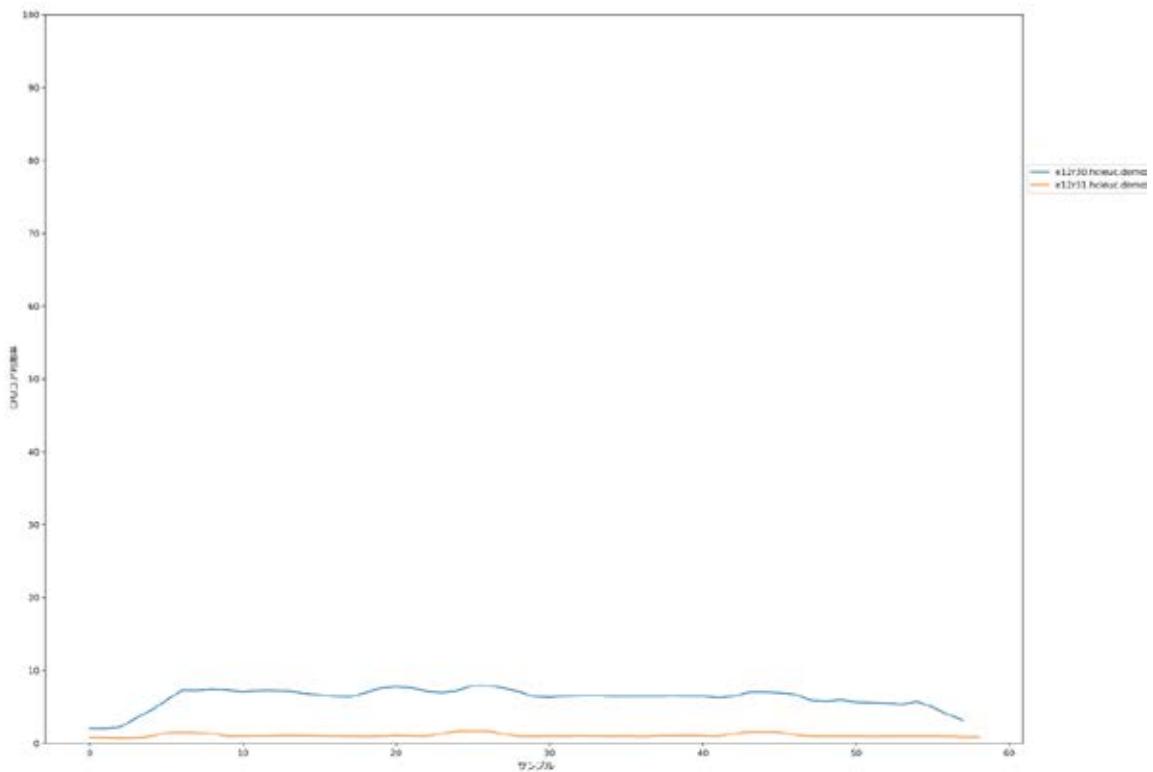
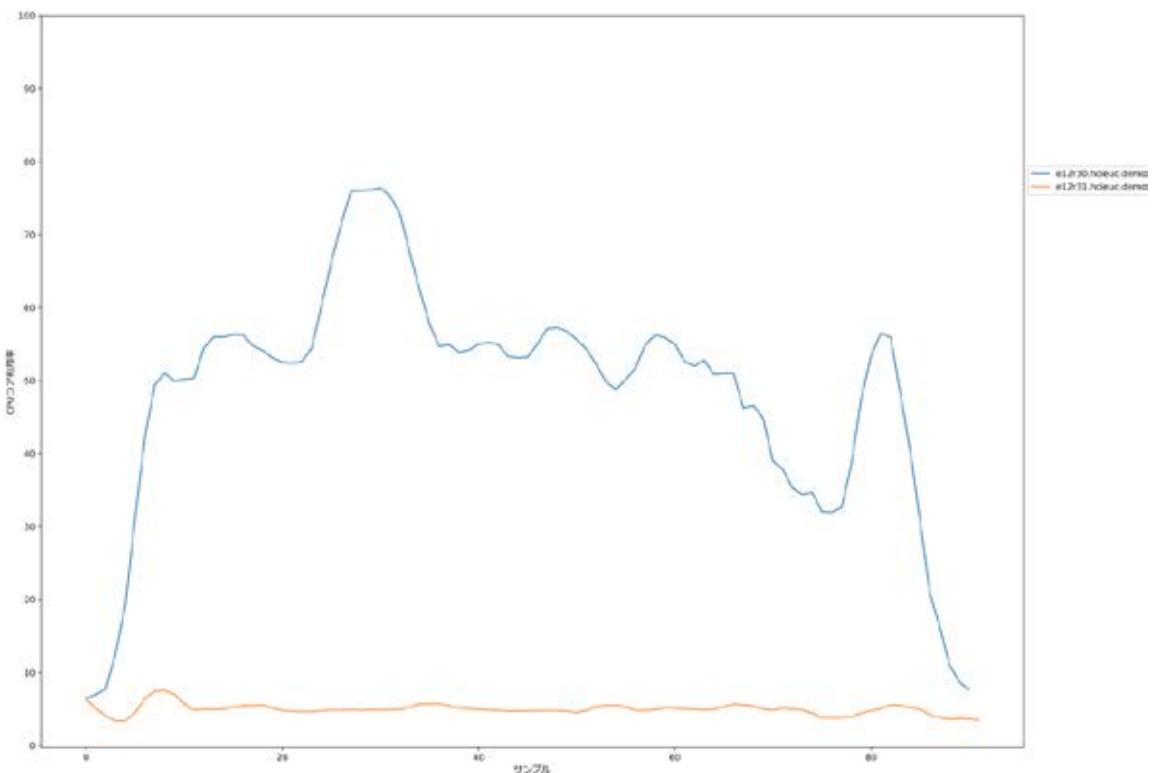


図 20) CATIA : vSphere の CPU 利用率 – 4Qx12



GPU 利用率は VM 1 台（図 21）および 12 台（図 22）ともに高い値で推移しました。

図 21) CATIA : GPU 利用率 – 1x16Q

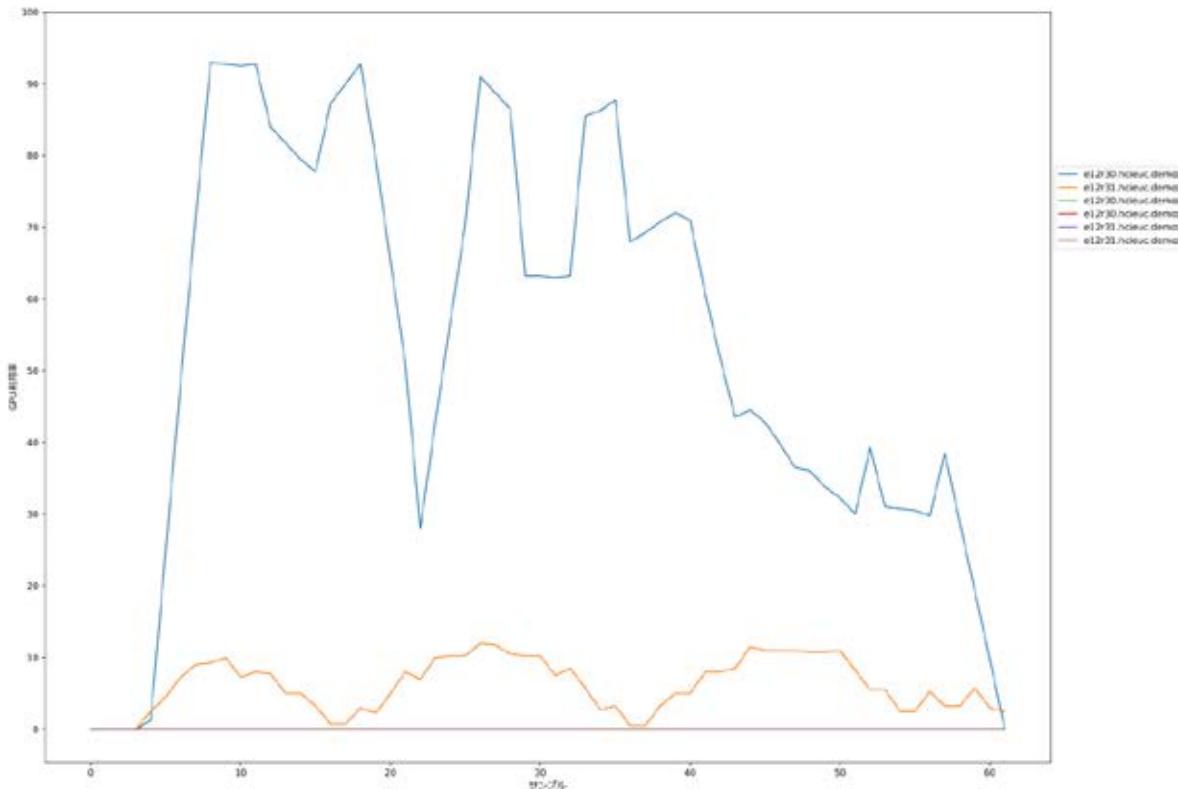


図 22) CATIA : GPU 利用率 – 12x4Q

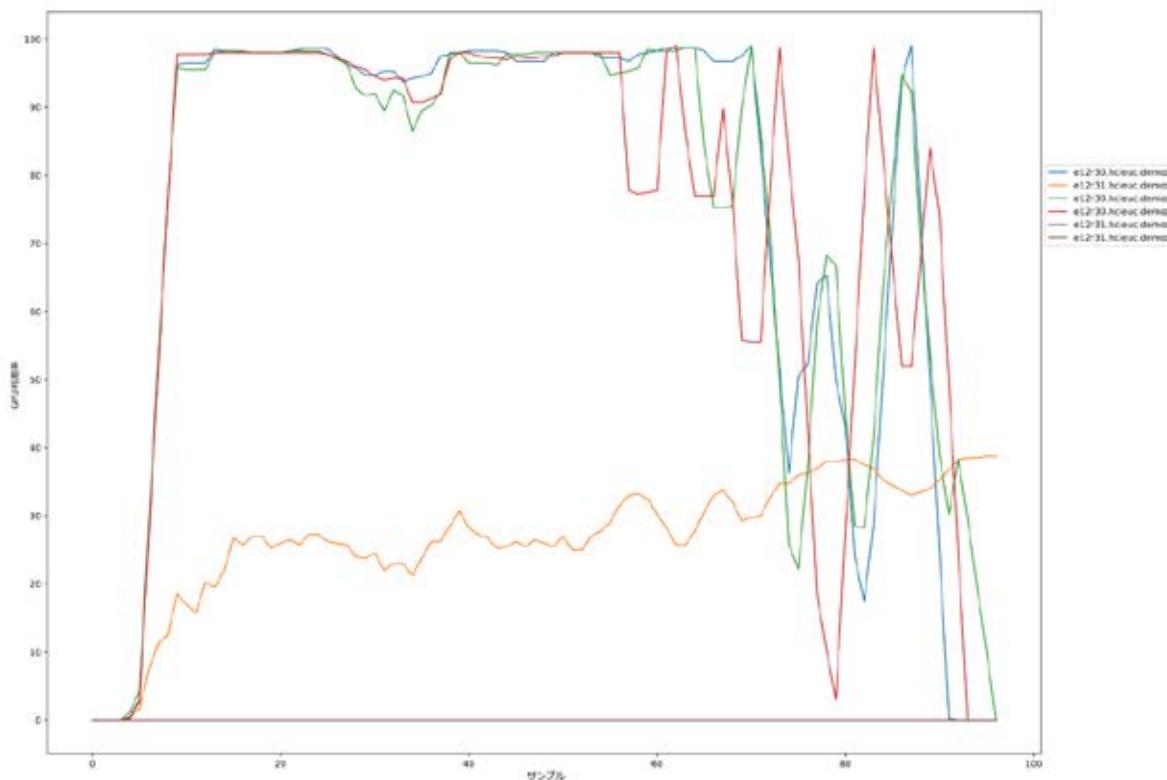
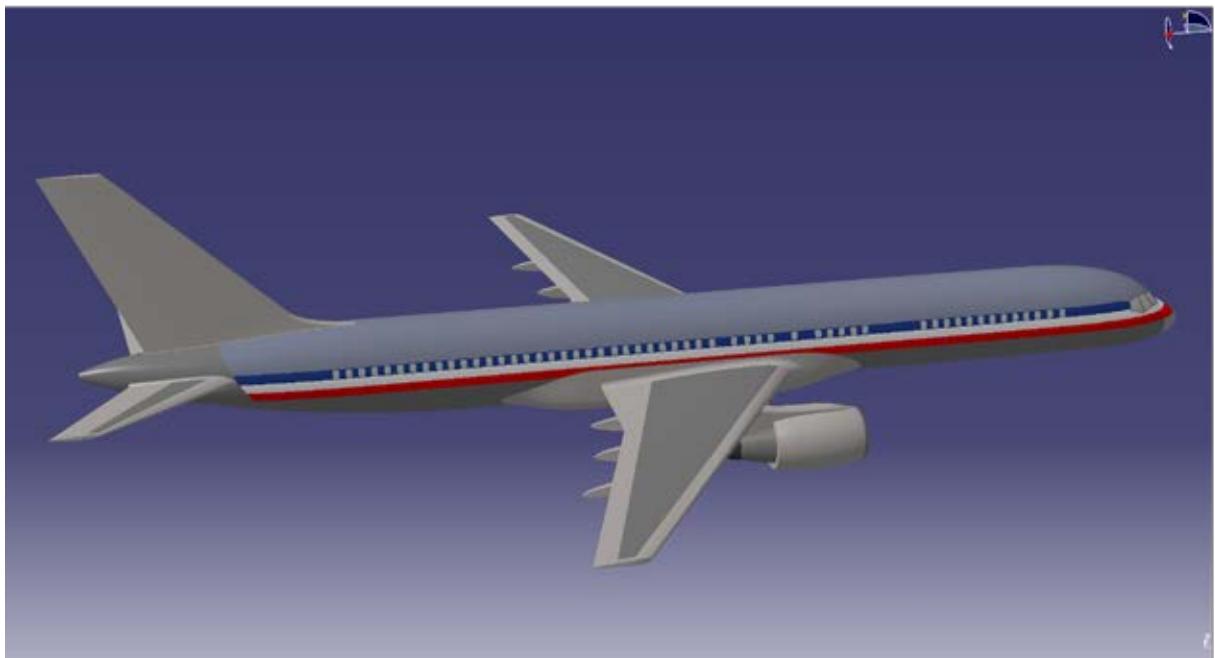


図 23 は、テスト中にキャプチャしたサンプルスクリーンショットです。

図 23) CATIA のサンプル



6.4 Creo (creo-02)

creo-02 ビューセットは、PTC の Creo 3 アプリケーションと Creo 4 アプリケーションで生成されたグラフィックス ワークロードのトレースから作成されました。モデル サイズは頂点数 2,000 万～4,800 万です。

このビューセットには、アプリケーションでサポートされる多数のレンダリング モードが含まれています。

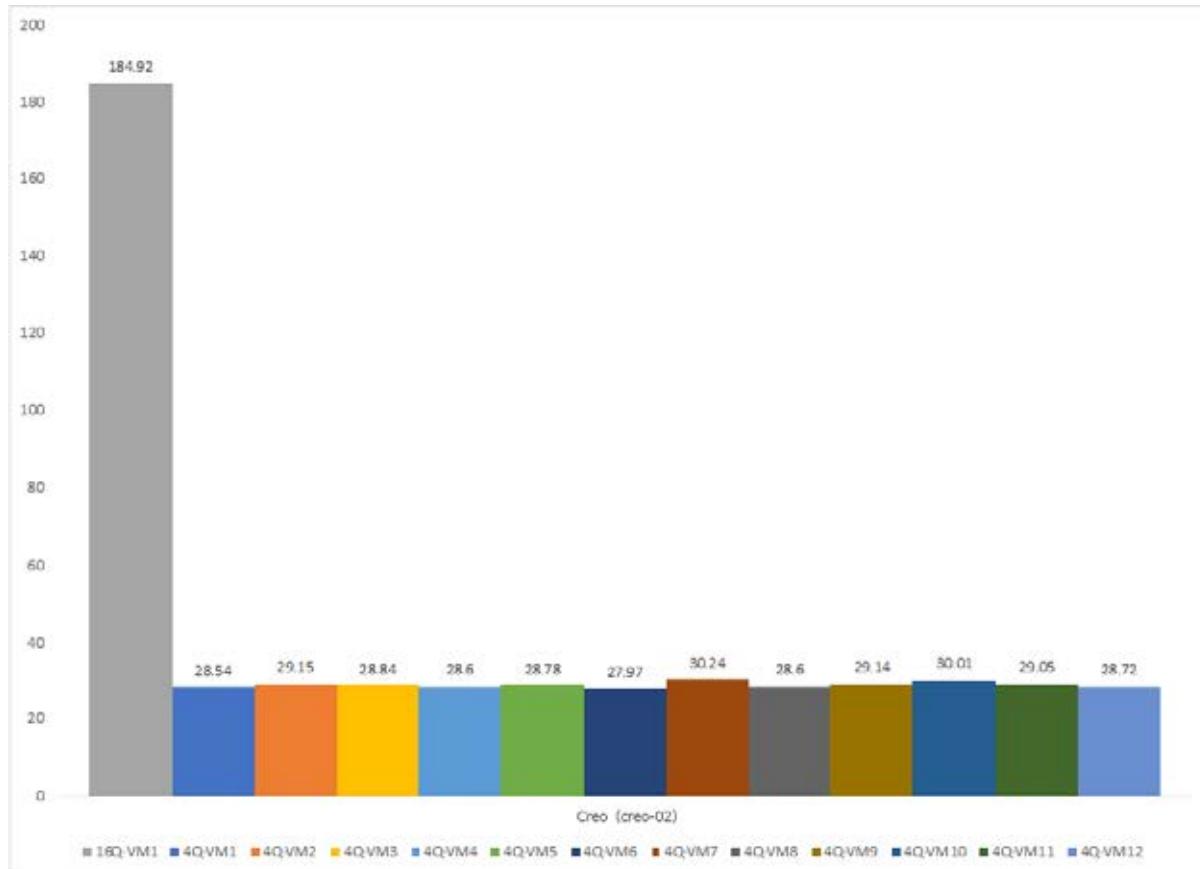
次のビューセット テストを実行しました。

- ワールドカー (シェーディング モード、環境マッピングによる反射、テクスチャ スペース バンプ マッピング、画像背景、スクリーン スペース アンビエント オクルージョンを適用)
- ワールドカー (シェーディング モード、反射、バンプ マッピング、画像背景、アンビエント オクルージョン、マルチサンプリング アンチエイリアス[4 倍]を適用)
- ワールドカー (シェーディング モード、反射、バンプ マッピング、画像背景、アンビエント オクルージョン、マルチサンプリング アンチエイリアス[8 倍]を適用)
- ワールドカー (シェーディング モード)
- エンジン (シェーディング モード)
- バイク (シェーディング モード、マルチサンプリング アンチエイリアス 4 倍を適用)
- ワールドカー (エッジ付きシェーディング モード、マルチサンプリング アンチエイリアス 4 倍を適用)
- エンジン (エッジ付きシェーディング モード)
- バイク (エッジ付きシェーディング モード)
- 爆撃機 x4 (エッジ付きシェーディング モード、マルチサンプリング アンチエイリアス 8 倍を適用、PTC Creo 4 からトレース)
- エンジン x4 (ワイヤフレーム モード、マルチサンプリング アンチエイリアス 4 倍を適用)
- 爆撃機 x4 (ワイヤフレーム モード、PTC Creo 4 からトレース)
- ワールドカー (隠線モード)
- バイク (隠線モード、マルチサンプリング アンチエイリアス 8 倍を適用)

- エンジン（隠線なしエッジ モード）
- 爆撃機 x4（隠線なしエッジ モード、マルチサンプリング アンチエイリアス 8 倍を適用、PTC Creo 4 からトレース）

16Q vGPU プロファイルの VM 1 台および 4Q vGPU プロファイルの VM 12 台の総合スコアを次に示します。

図 24) Creo の総合スコア



Creo ビューセットでは、VM 12 台の vSphere のコア利用率は 70% 以下です（図 26）。

図 25) Creo : vSphere の CPU 利用率 – 16Qx1

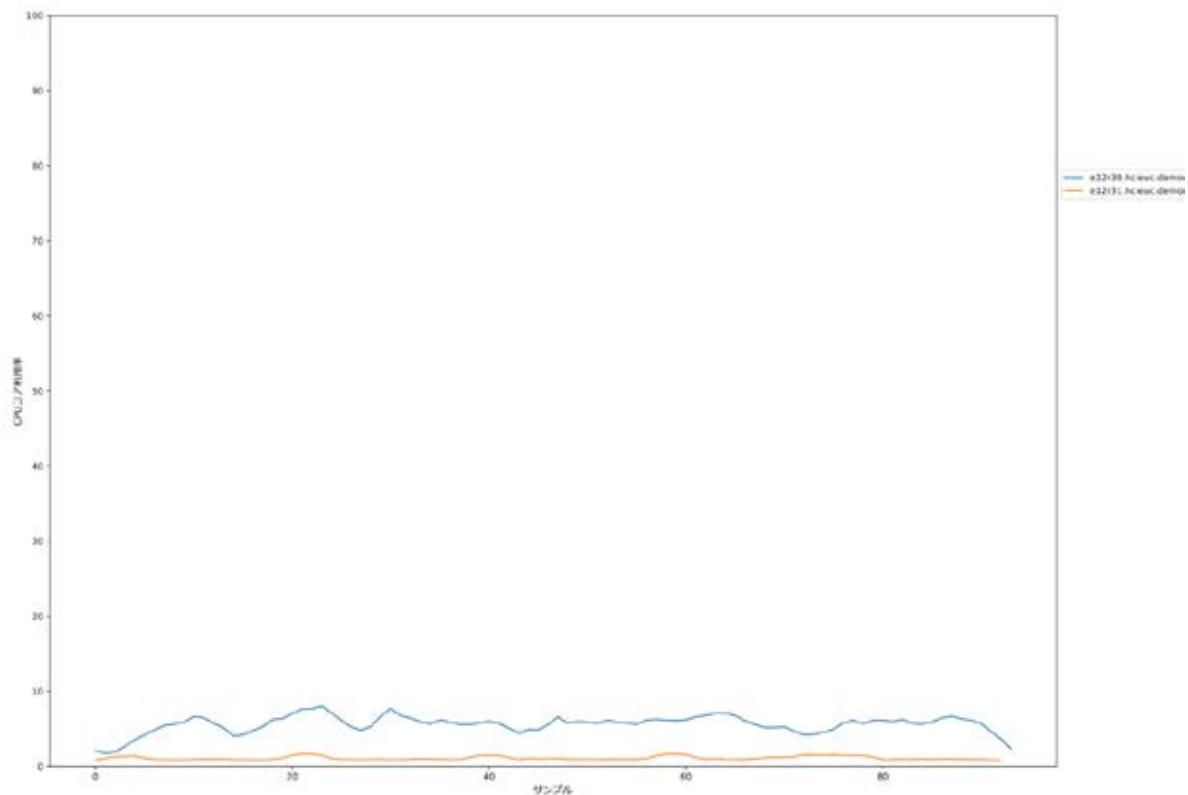
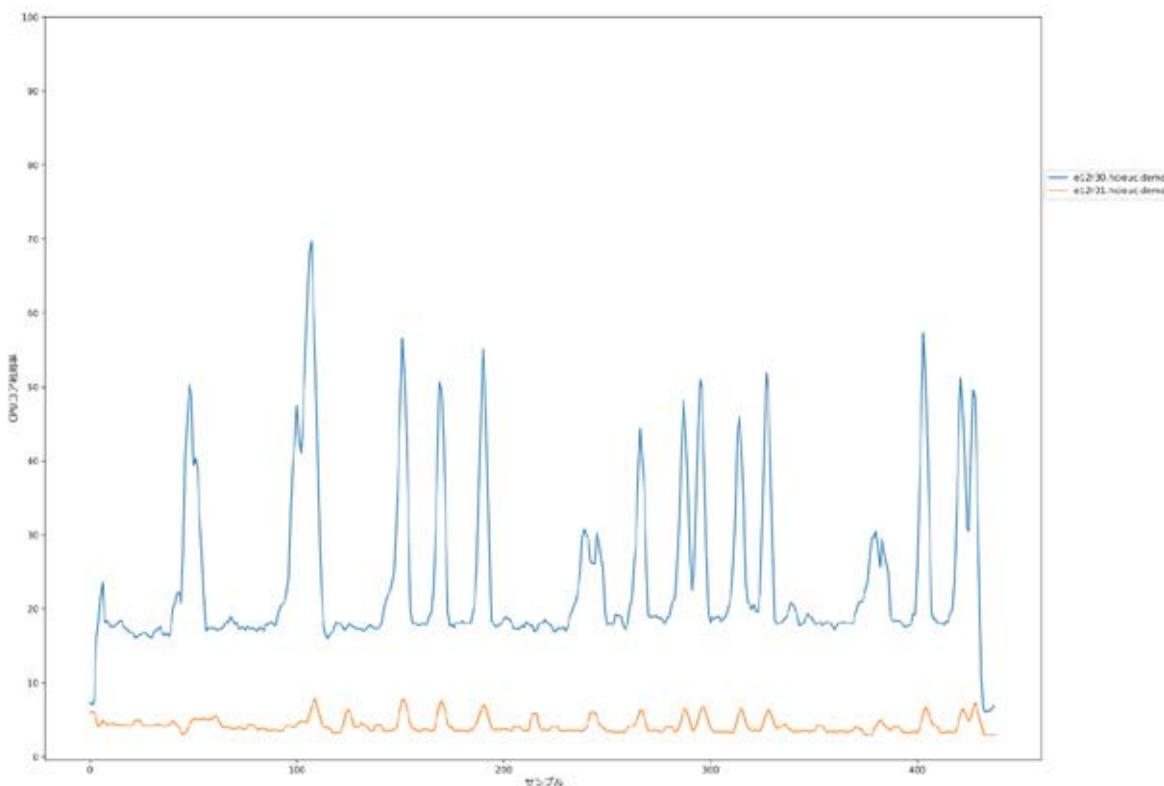


図 26) Creo : vSphere の CPU 利用率 – 4Qx12



VM 1 台の場合の Creo ビューセットの GPU 利用率を図 27 に示します。VM 12 台の場合の結果を図 28 に示します。

図 27) Creo : GPU 利用率 – 16Qx1

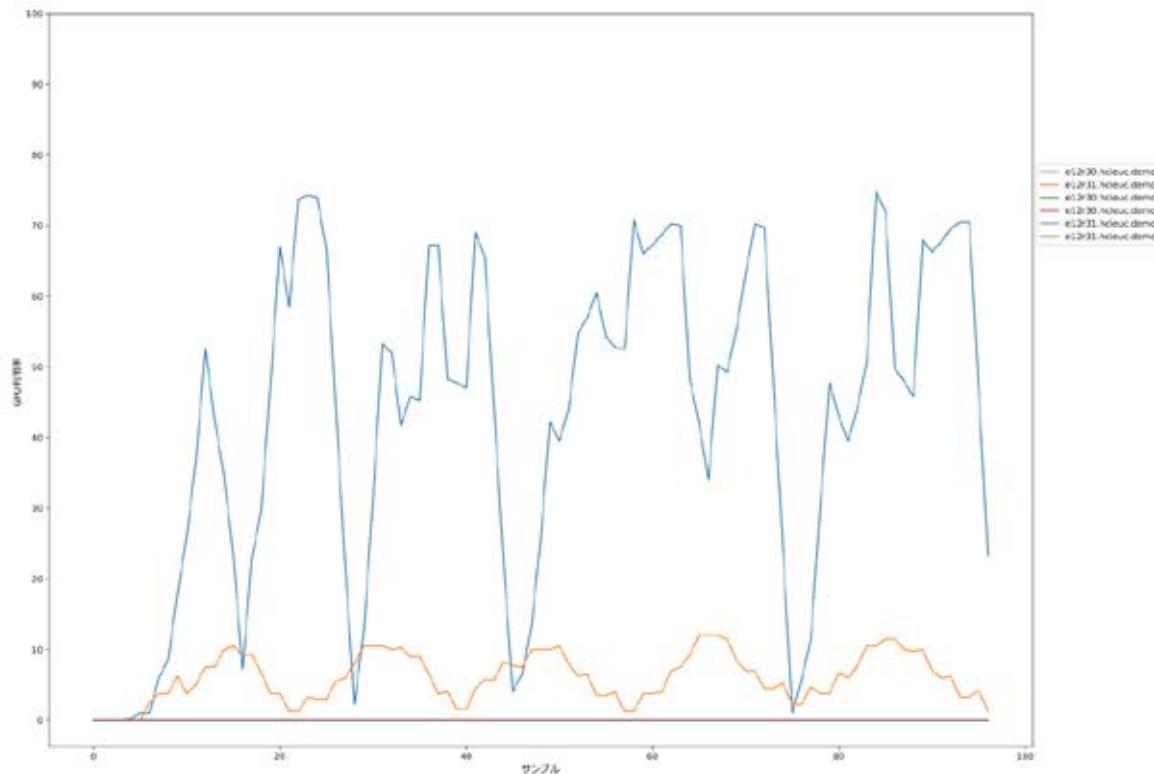


図 28) Creo : GPU 利用率 – 4Qx12

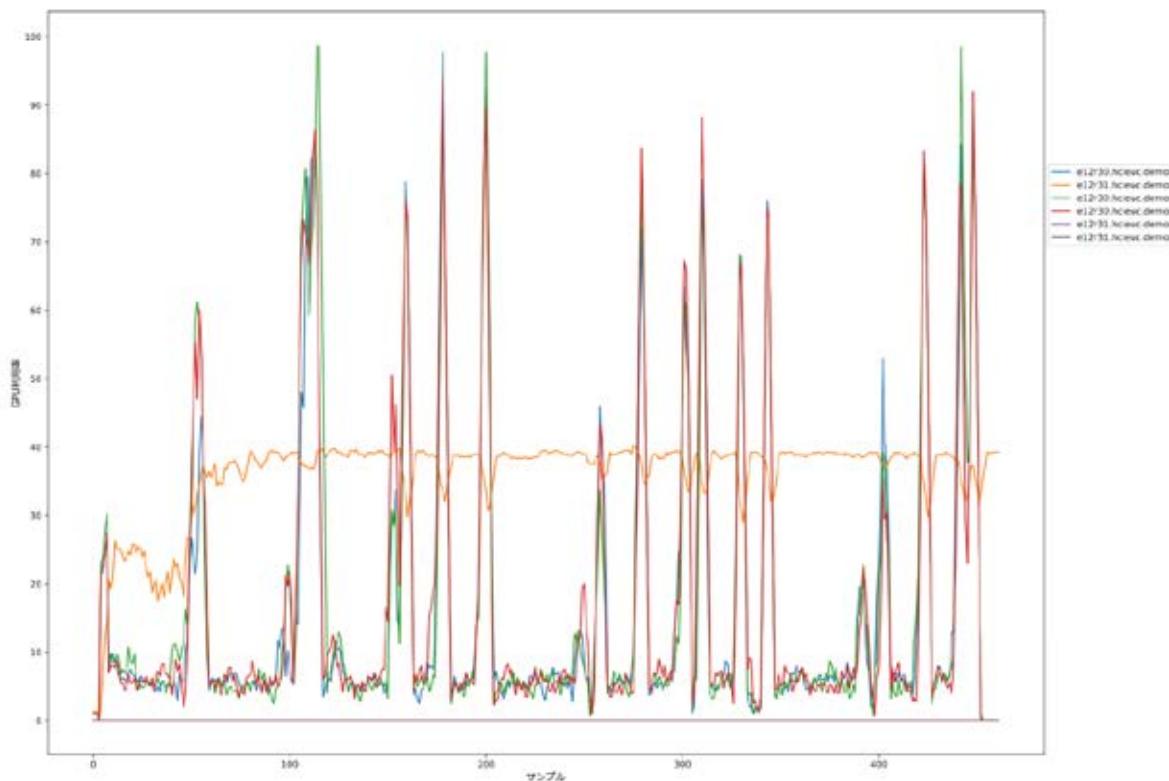
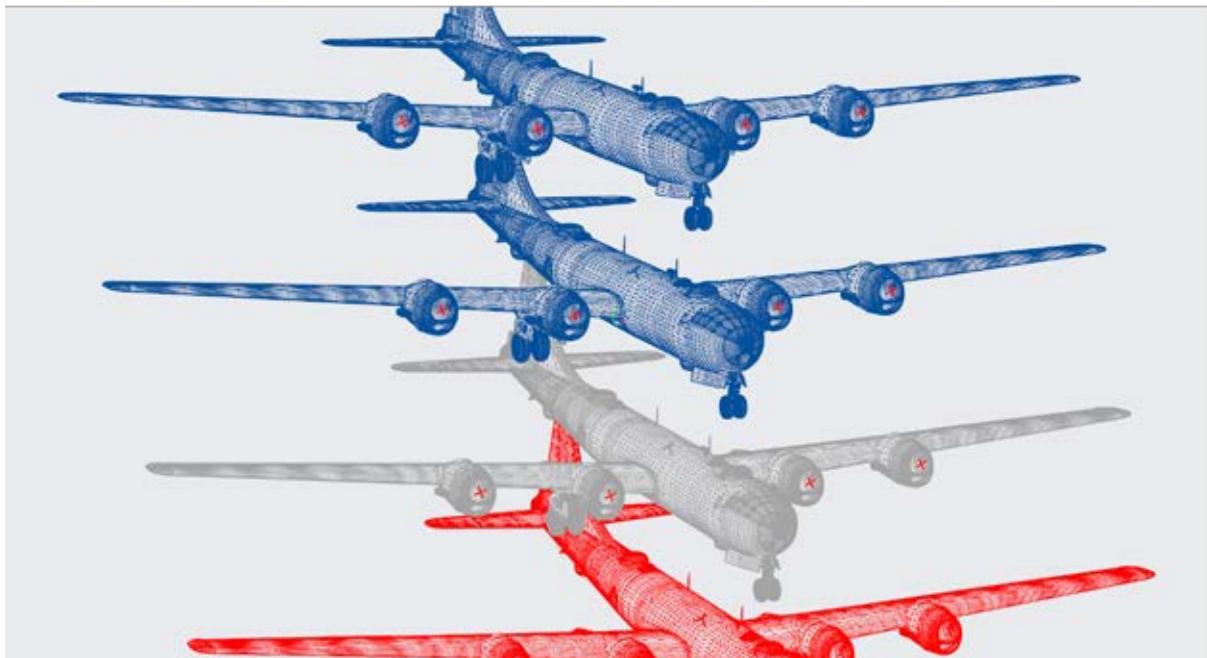


図 29 は、実施されたテストからキャプチャしたサンプルスクリーンショットです。

図 29) Creo のサンプル



6.5 Energy (energy-02)

energy-02 ビューセットは、オープンソースの地震可視化アプリケーション OpendTect で使用されているレンダリング手法に基づいています。MRI や CT などの医療用画像と同様、物理探査でもサブサーフェスを使用して画像スライスを生成し、3D グリッドに組み込みます。ボリュームレンダリングでは、この 3D ボリュームトリック グリッドを 2D に投影することで、さらに詳細な分析と解釈を行います。

すべてのフレームで、ボリュームの境界のキューブ面がテッセレーションされ、フラグメントシェーダでレンダリングされます。このシェーダは、視点からボリュームまでのレイキャスティングを実行します。ピクセルが完全に不透明になるか、光線がボリュームを抜けるまで、透明光源のカラーマッピング値を積算します。

3D グリッドのボクセルは単一のスカラー値です。伝達関数 (1D ルックアップ テーブル) は、3D 密度値をカラー値とアルファ値にマッピングします。ライティングの計算では、各ボクセルの中心差分を使用して勾配がオンザフライで算出されます。このような状態の変化は、グラフィックスサブシステムのさまざまな要素をシミュレートします。このビューセットは、3D テクスチャ (つまりトリアニニア補間) のハードウェアサポートを利用します。

ボリュームレンダリングに加えて、テストにはインライン面とクロスライン面 (X 面と Y 面のスライス) の両方が含まれます。また、一部のサブテストには「層位」が存在します。層位は興味対象となる地層境界で、物理探査学者が生成し、テクスチャ付きトライアングルストリップを使用してレンダリングされます。

このビューセットで使用されている 3D データセットは実際の地震データセットで、[SEG Wiki の Open Data のページ](#)から入手できます。データセットはオリジナルの SEG-Y 形式から変換され、JPEG-2000 を使用して圧縮されました。

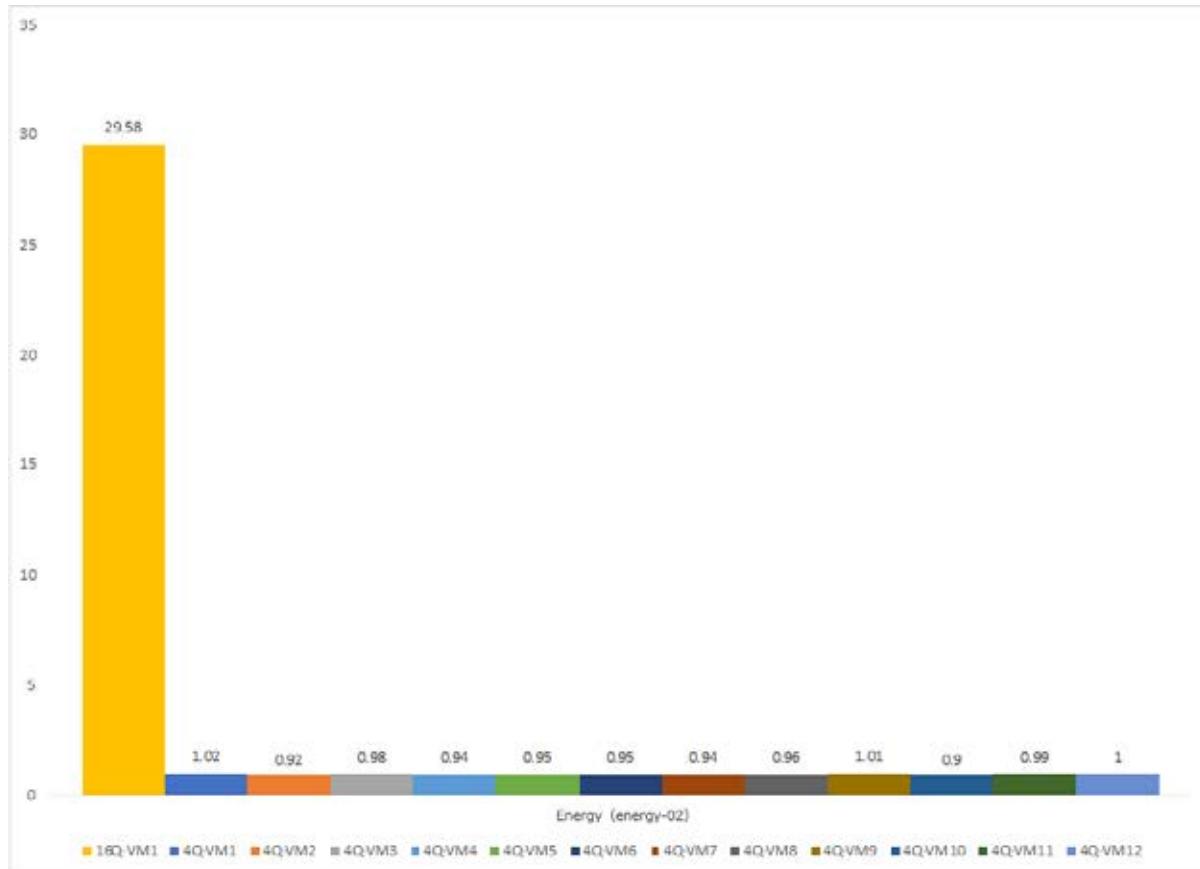
次のビューセット テストを実行しました。

- ブレイク リッジ ボリューム (1307x95x1300) と層位
- F3 オランダ ボリューム (950x450x462) と層位
- オプナキ ボリューム (1949x731x1130)
- ブレイク リッジ ボリューム (アニメーションクリップ面あり) と層位

- F3 オランダ ボリューム (アニメーション クリップ面あり) と層位
- オプナキ ボリューム (アニメーション クリップ面あり)

16Q VM 1 台と 4Q VM 12 台の総合スコアを図 30 に示します。フレーム バッファ サイズは、Energy ビューセットの総合スコアに大きく影響します。

図 30) Energy の総合スコア



Energy ビューセットでは、VM 12 台の vSphere の CPU 利用率は 80% 以下です（図 32）。

図 31) Energy : vSphere の CPU 利用率 – 16Qx1

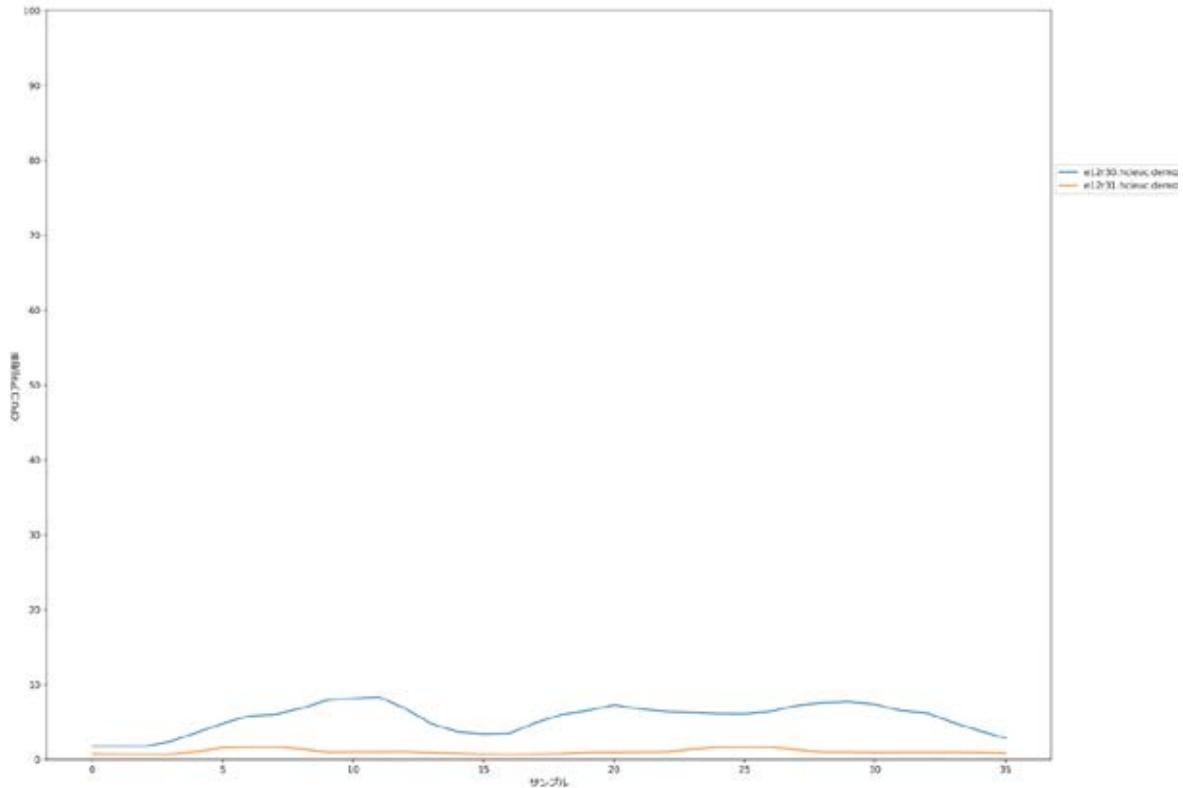


図 32) Energy : vSphere の CPU 利用率 – 4Qx12

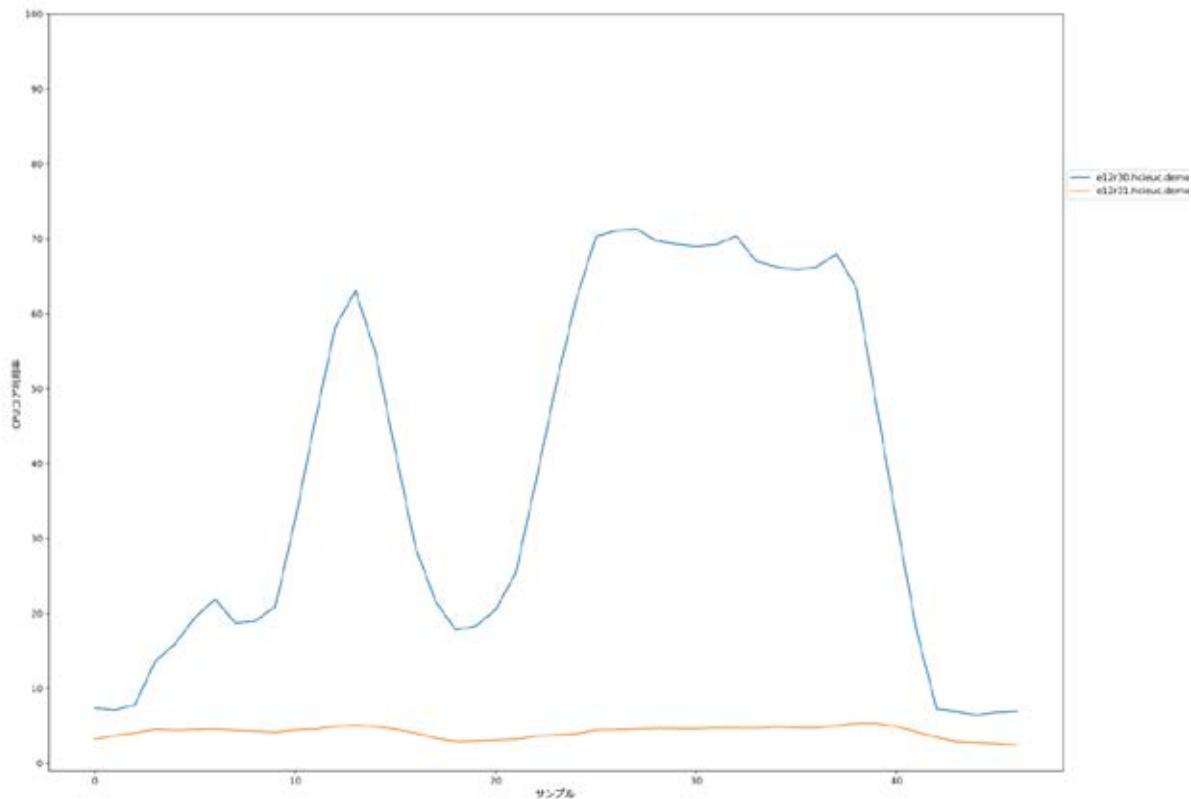


図 33) Energy : GPU 利用率 – 16Qx1

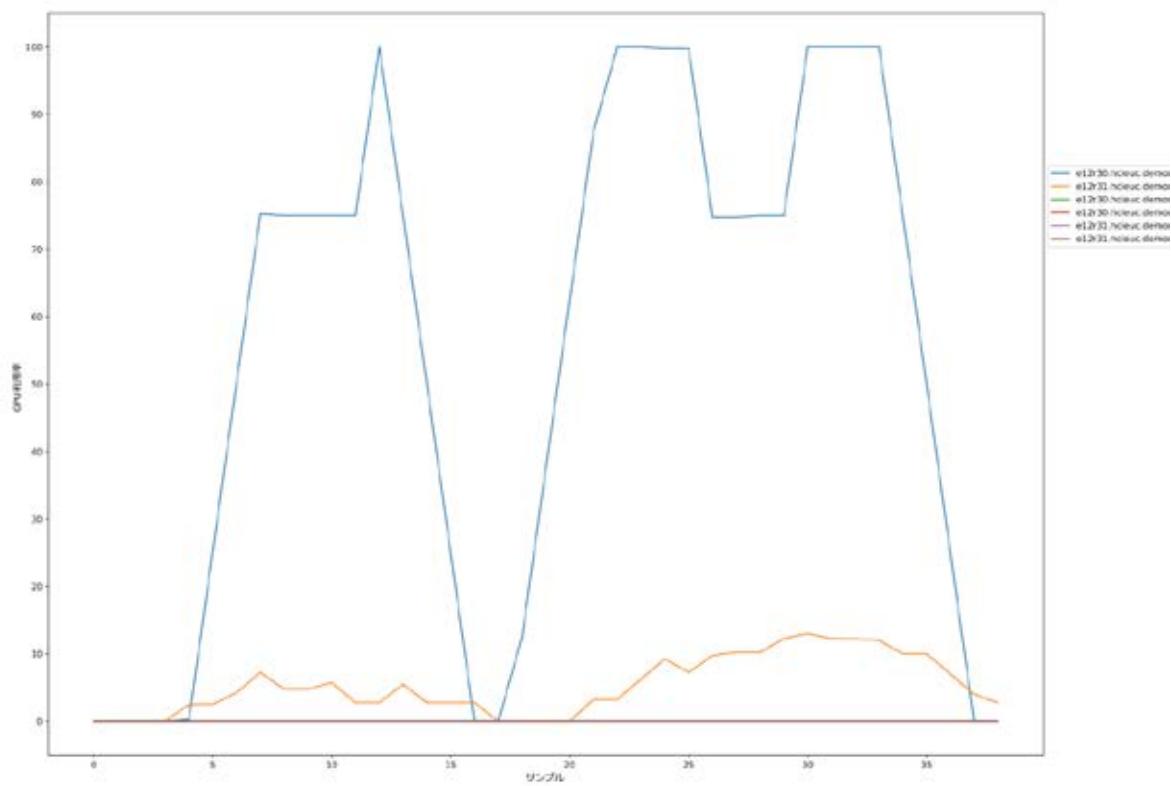


図 34) Energy : GPU 利用率 – 4Qx12

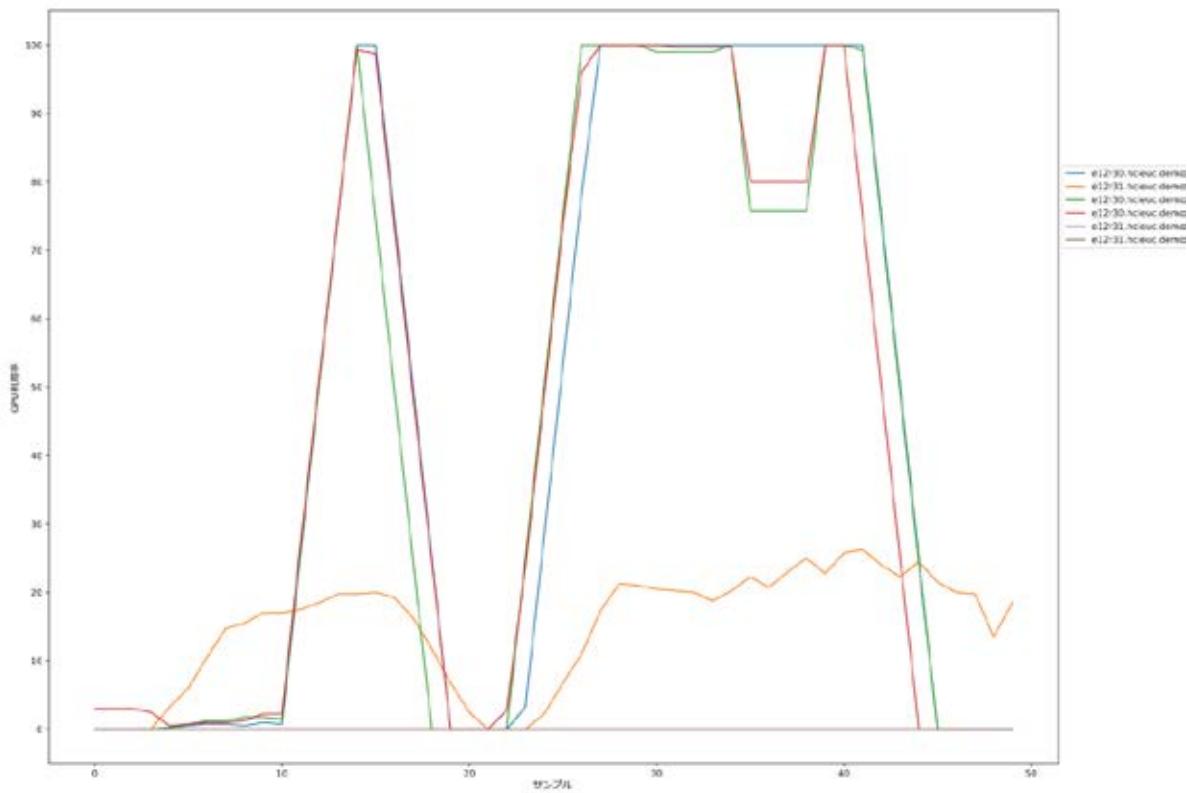
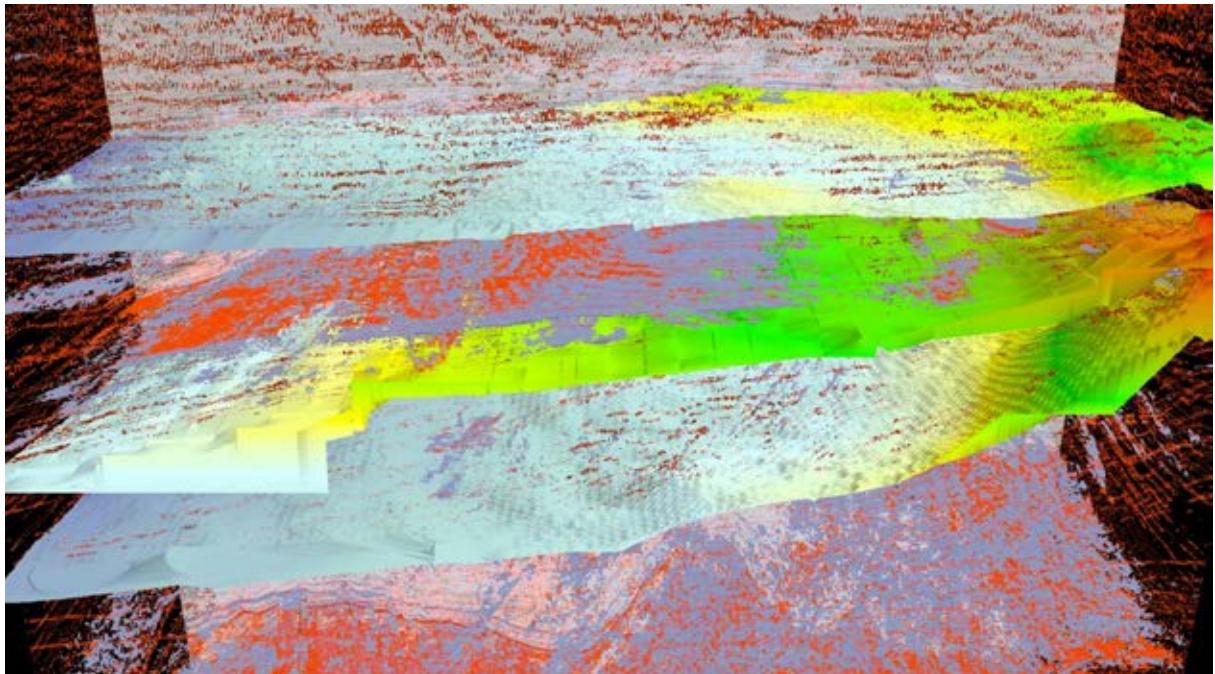


図 35 は、実施されたテストからキャプチャしたサンプルスクリーンショットです。

図 35) Energy のサンプル



6.6 Maya (maya-05)

maya-05 ビューセットは、Autodesk の Maya 2017 アプリケーションで生成されたグラフィックス ワークロードのトレースから作成されました。

このビューセットには、シェーディング モード、アンビエント オクルージョン、マルチサンプリング アンチエイリアス、透明度など、アプリケーションでサポートされる多数のレンダリング モードが含まれています。テストはすべて Viewport 2.0 を使用してレンダリングしました。

次のビューセット テストを実行しました。

- 玩具店（スムーズシェーディング[シェーディング上にワイヤフレームを適用]モード、アンビエント オクルージョン、マルチサンプリング アンチエイリアス[4 倍]）
- 玩具店（ワイヤフレーム モード、マルチサンプリング アンチエイリアス[8 倍]）
- ジャングル エスケープ（スムーズシェーディング[ハードウェア テクスチャを適用]モード、アンビエント オクルージョン）
- ジャングル エスケープ（スムーズシェーディング[ハードウェア テクスチャを適用]モード）
- スヴェン スペース（スムーズシェーディング[ハードウェア テクスチャを適用]モード）
- スヴェン スペース（スムーズシェーディング、アンビエント オクルージョン、マルチサンプリング アンチエイリアス 4 倍）
- HSM サテライト（スムーズシェーディング、マルチサンプリング アンチエイリアス[8 倍]）
- 船の水しぶき（スムーズシェーディング[すべてのライトを適用]）
- 船の水しぶき（ワイヤフレーム モード、マルチサンプリング アンチエイリアス 4 倍）
- 船の水しぶき（スムーズシェーディング[ハードウェア テクスチャを適用]モード、アンビエント オクルージョン、マルチサンプリング アンチエイリアス 8 倍）

16Q VM 1 台と 4Q VM 12 台を使用したテストの総合スコアを図 36 に示します。

図 36) Maya の総合スコア

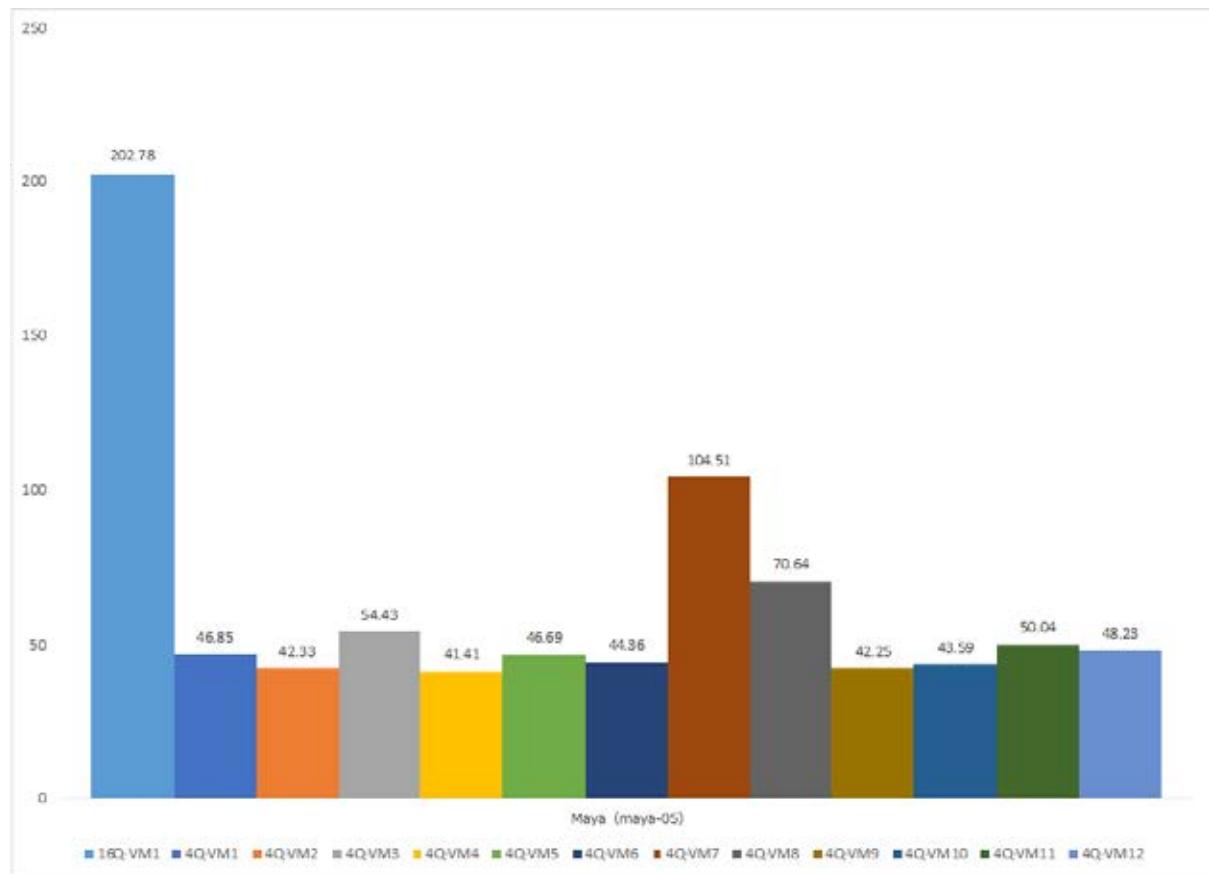


図 37) Maya : vSphere の CPU 利用率 – 16Qx1

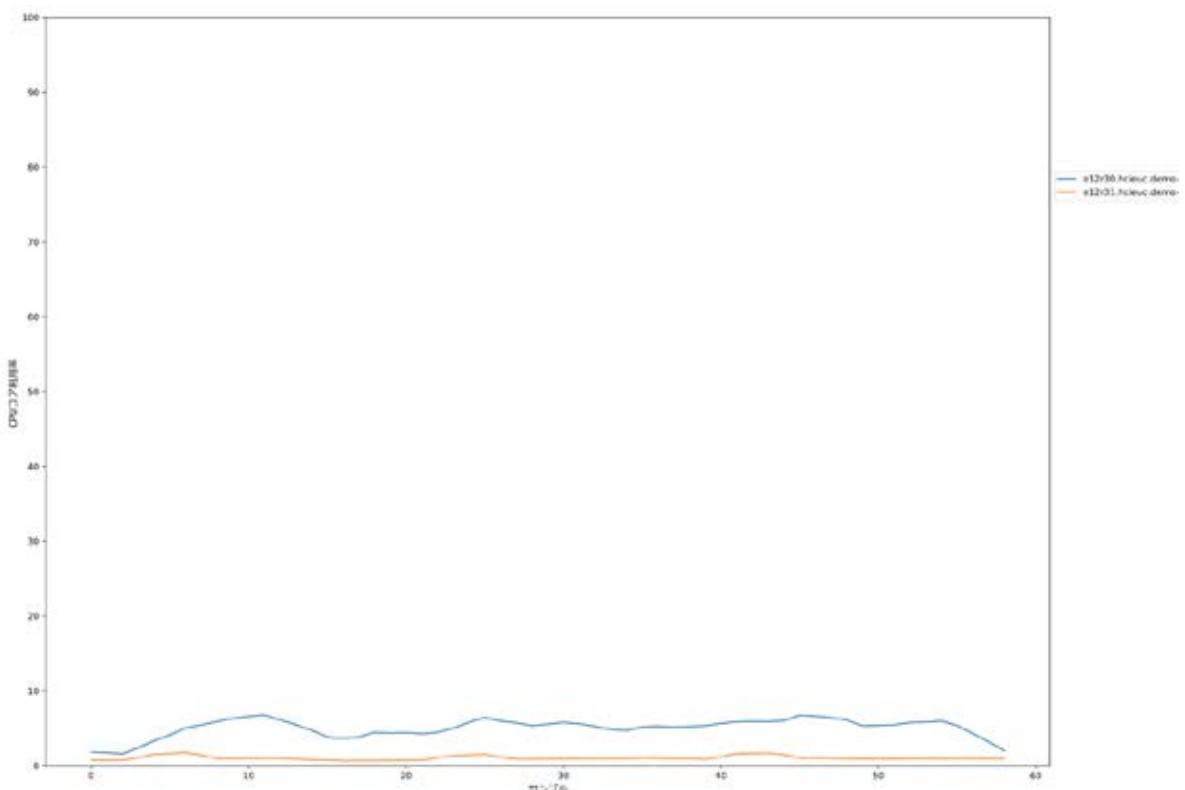


図 38) Maya : vSphere の CPU 利用率 – 4Qx12

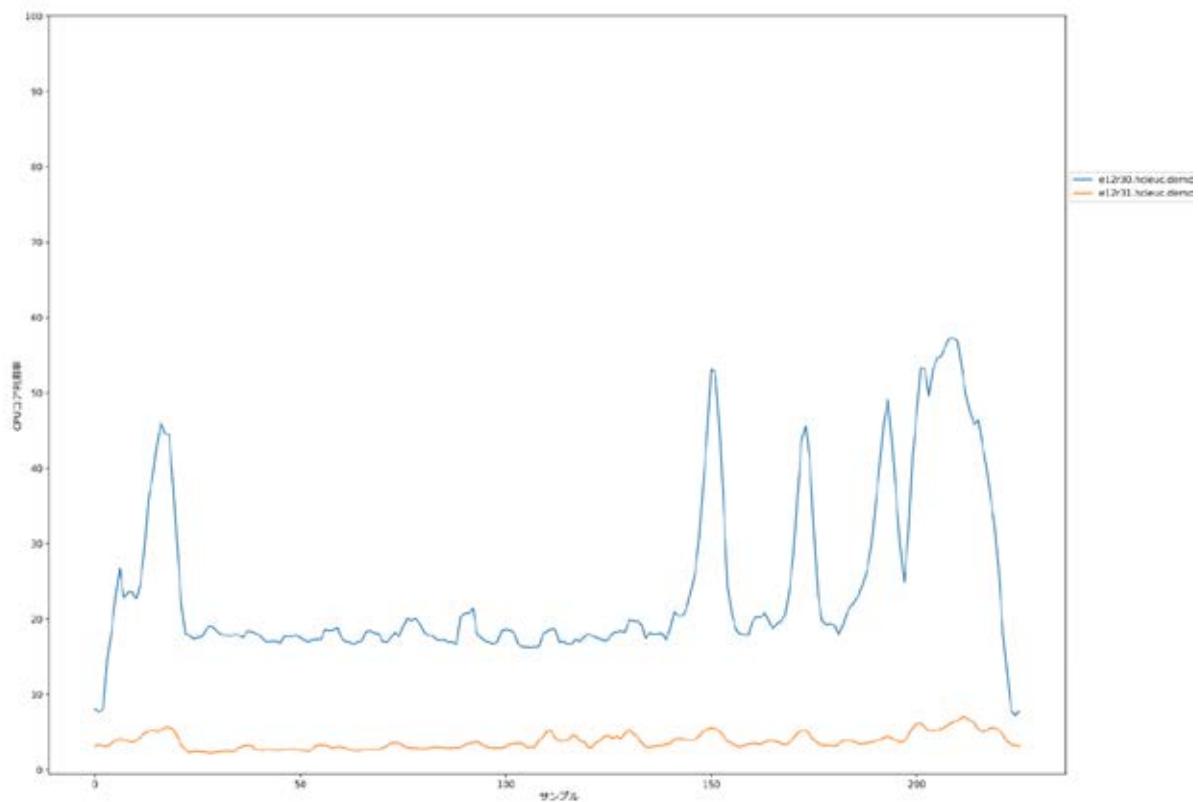


図 39) Maya : GPU 利用率 – 16Qx1

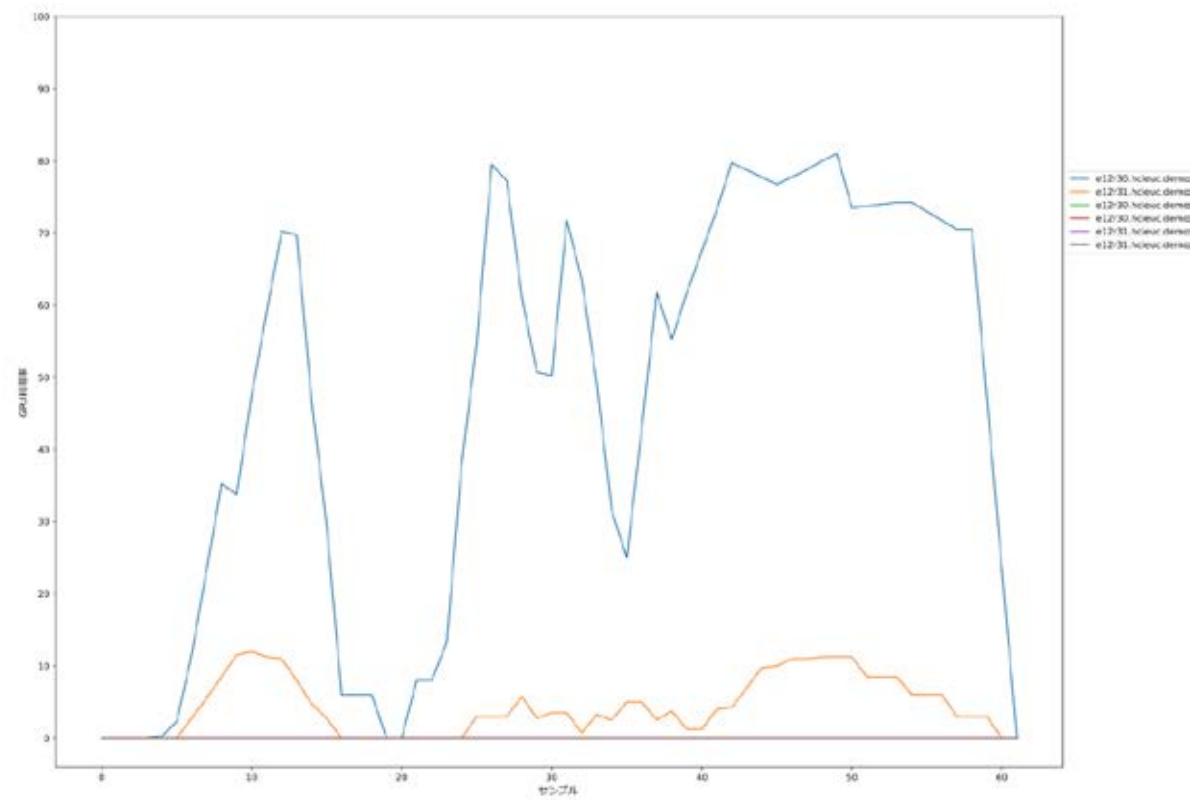


図 40) Maya : GPU 利用率 – 4Qx12

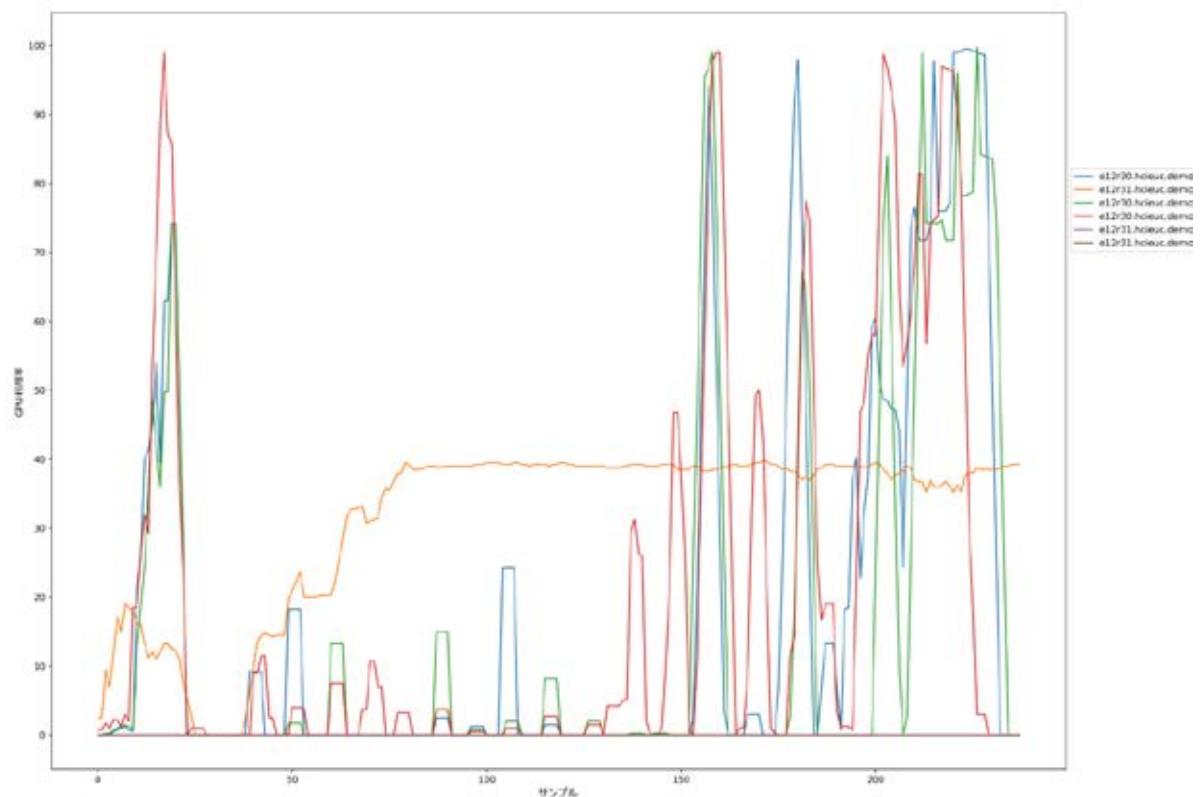


図 41) Maya のサンプル



6.7 Medical (medical-02)

medical-02 ビューセットでは、[ImageVis3D ボリューム ビジュアライゼーション プログラム](#)の Tuvok レンダリング コアを使用します。Tuvok は 3D ボリュームトリック グリッドの 2D 投影をレンダリングします。このビューセットの典型的な 3D グリッドは、スキャナ（CT や MRI など）で取得された一連の 3D スライスです。

2 つのレンダリング モード（スライスベースのレンダリングとレイキャスティング）を使用しました。

スライスベースのレンダリングでは、現在の視野角に揃えられた同一平面上の一連のスライスが CPU で計算されてグラフィックス ハードウェアに送信され、テクスチャリングやその他の計算が実施されます。計算には、内部構造を明らかにするための伝達関数のルックアップ、ライティング、クリッピングが含まれます。最後に、スライスがブレンドされて画像が表示されます。

レイキャスティングでは、ボリュームまで光線を送り、完全に不透明になるかまたは光線がボリュームを抜けるまで、透明光源のカラーマッピング値を積算します。

スライスベースとレイキャスト両方のレンダリングで、ボリュームはさらに 512x512x512 の 3D ボリュームに分割されます。この手法はブリッキングと呼ばれ、一般に多くの GPU ハードウェアでレンダリング パフォーマンスが向上します。

3D グリッドのボクセルは単一のスカラー値です。伝達関数（1D または 2D ルックアップ テーブル）は、3D 密度値をカラー値とアルファ値にマッピングします。2D テーブルの場合、第 2 軸は各サンプルにおける勾配強度として定義されます。ライティングの計算では、各ボクセルの中心差分を使用して勾配がオンザフライで算出されます。このような状態の変化は、グラフィックス サブシステムのさまざまな要素をシミュレートします。このビューセットは、3D テクスチャ（つまりトリニア補間）のハードウェア サポートを利用します。

このビューセットには次の 4 つのデータセットがあります。

- 経時に反復される複数の 3D ボリュームで構成される 4D 心臓データセットで、位相コントラスト MRI スキャナから取得されました。データセットは 80MB あり、スタンフォード大学医学部とルシール・パッカード小児病院の放射線科から提供されたものです。各ボリュームは 256x256x32 の 16 ビット サンプルで構成されています。
- ウィーン工科大学から提供されたクワガタムシ データセット。データセットのサイズは 650MB で、大容量のメモリを必要とするワークロードです。ボリュームは 832x832x494 の 16 ビット サンプルで構成されています。
- SPECgpc 委員会メンバー 1 名の頭部 MRI スキャン。このメンバーは、SPECviewperf で使用するためにこのデータを公開しました。ボリュームは 232x256x192 の 16 ビット サンプルで構成されています。
- 同じ SPECgpc 委員会メンバーの右上部胸郭と腕の CT スキャン。SPECviewperf で使用するためにこのデータも公開しました。ボリュームは 512x512x102 の 16 ビット サンプルで構成されています。

このビューセットのテストは前述の 4 つのデータセットから次のように生成されました。

- 4D 心臓、1D 伝達関数、スライスベースのレンダリング
- 4D 心臓、1D 伝達関数、レイキャスティング
- クワガタムシ、1D 伝達関数、スライスベースのレンダリング
- クワガタムシ、1D 伝達関数、レイキャスティング
- 頭部 MRI、2D 伝達、レイキャスティング
- 頭部 MRI、2D 伝達、レイキャスティング、クリップ面
- 胸郭 CT、2D 伝達、レイキャスティング
- 胸郭 CT、2D 伝達、レイキャスティング、クリップ面

[Tuvok レンダリング コア](#)には、MIT のオープンソース ライセンスが必要です。Tuvok にはヒルベルト曲線（ライス大学が 1998 年に著作権を取得）が実装されています。また、2 条項 BSD ライセンスが必要な LZ4 も含まれています。

図 42) Medical の総合スコア

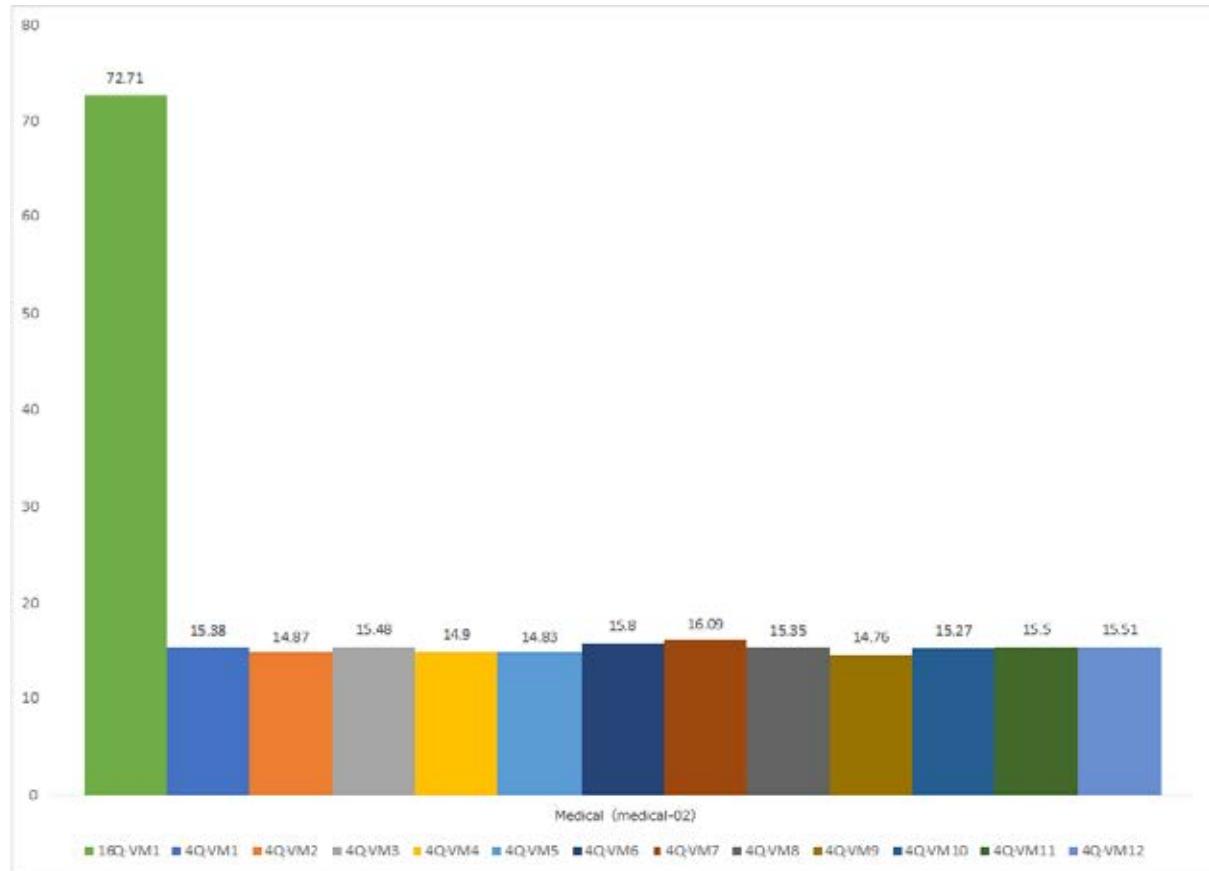


図 43) Medical : vSphere の CPU 利用率 – 16Qx1

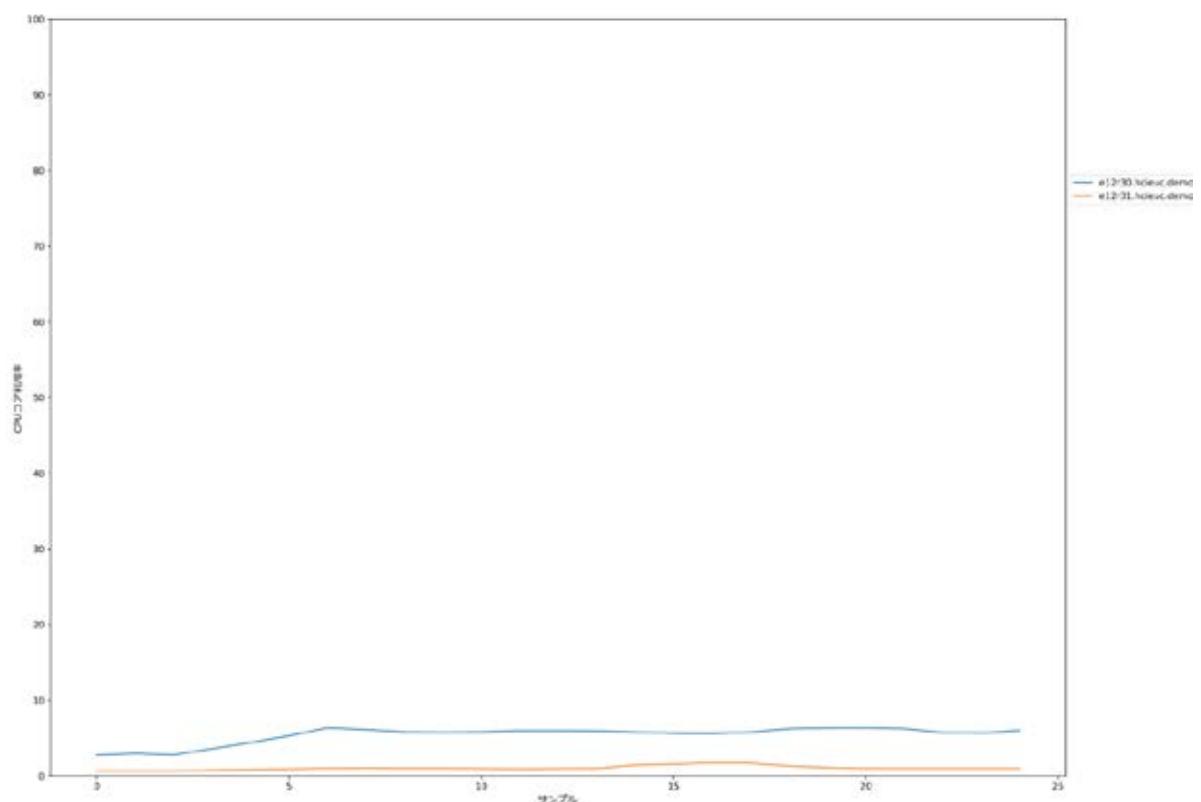


図 44) Medical : vSphere の CPU 利用率 – 4Qx12

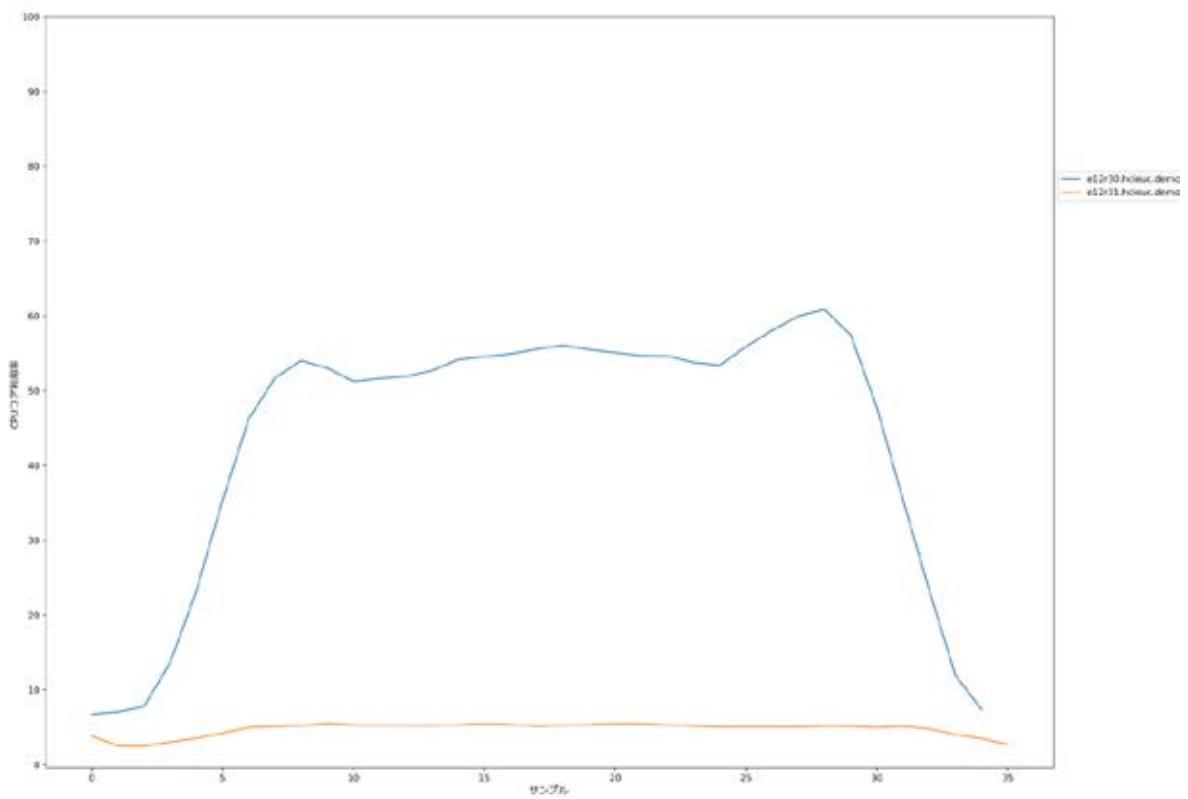


図 45) Medical : GPU 利用率 – 16Qx1

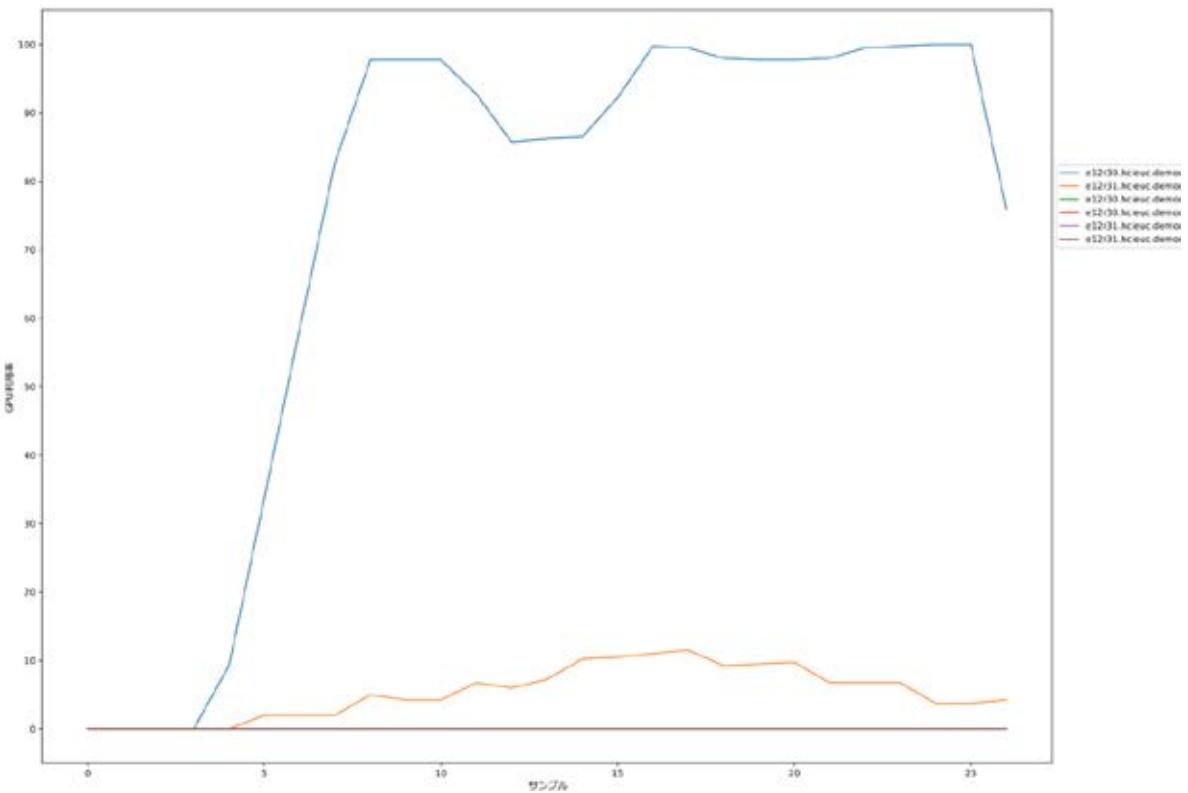


図 46) Medical : GPU 利用率 – 4Qx12

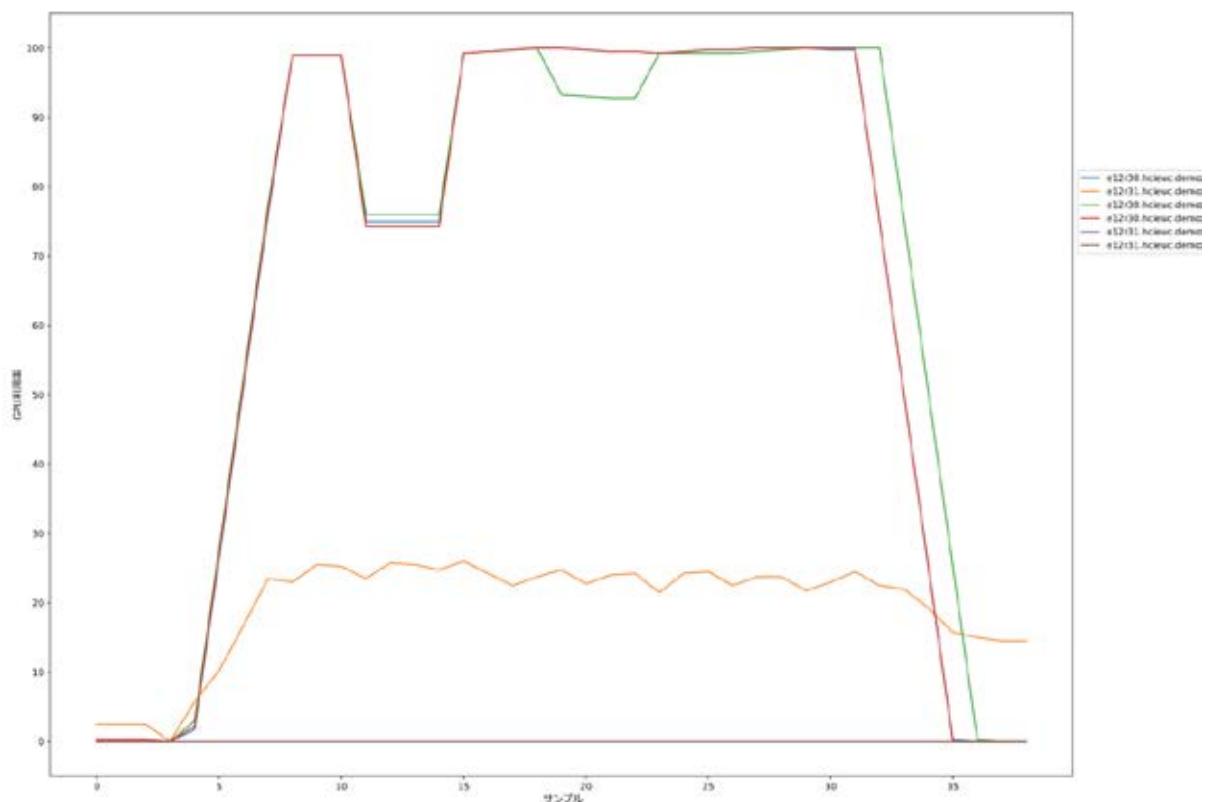
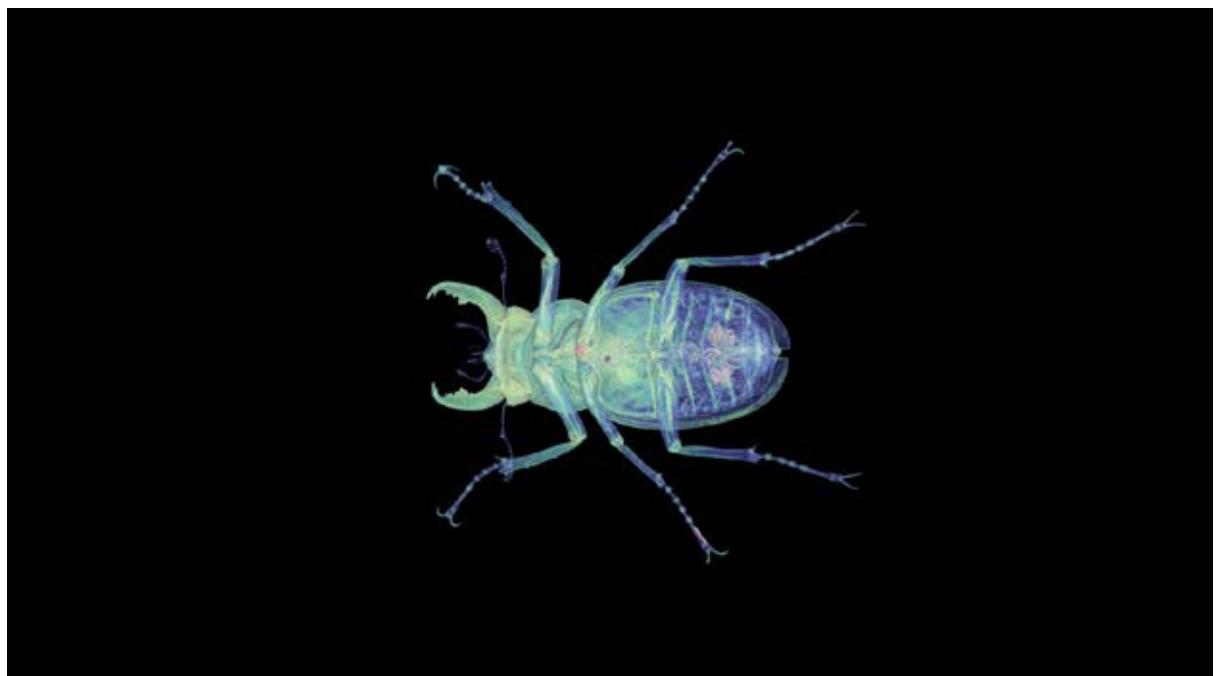


図 47) Medical のサンプル



6.8 Showcase (showcase-02)

showcase-02 ビューセットは、Autodesk の Showcase 2013 アプリケーションのトレースから作成されました。このビューセットで使用されているモデルの頂点数は 800 万です。

このビューセットは DX レンダリングを使用しています。レンダリング モードは、シェーディング、投影シャドウ、およびセルフシャドウです。

このビューセットには次のテストが含まれます。

- シェーディング (セルフシャドウを適用)
- シェーディング (セルフシャドウと投影シャドウを適用)
- シェーディング
- シェーディング (投影シャドウを適用)

図 48) Showcase の総合スコア

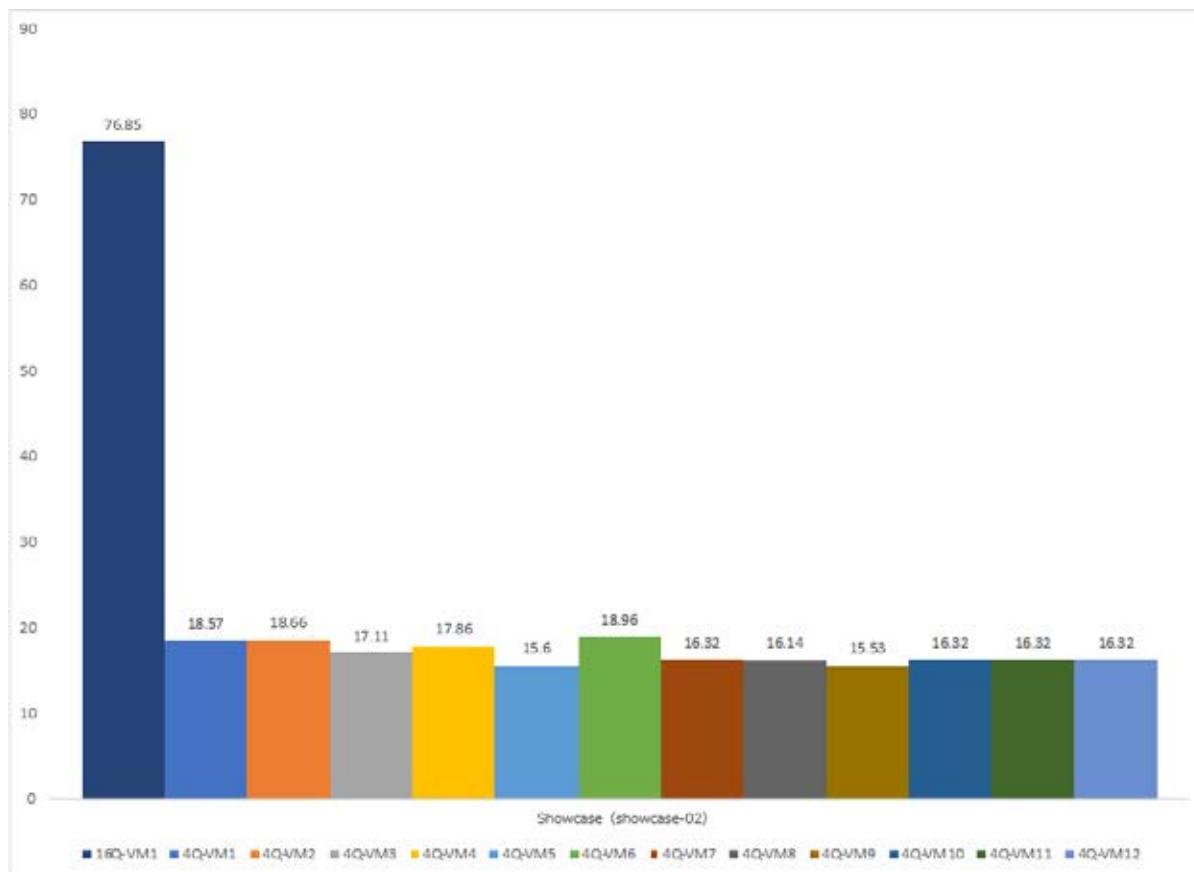


図 49) Showcase : vSphere の CPU 利用率 – 16Qx1

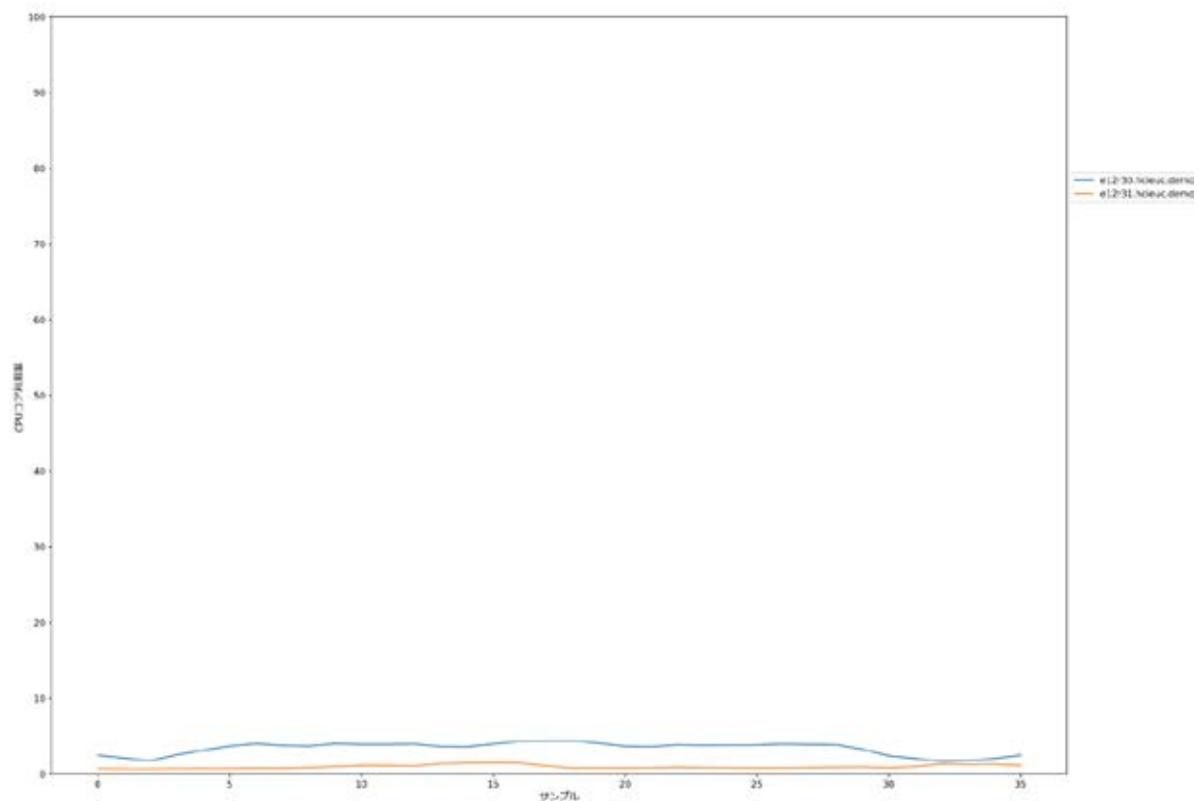


図 50) Showcase : vSphere の CPU 利用率 – 4Qx12

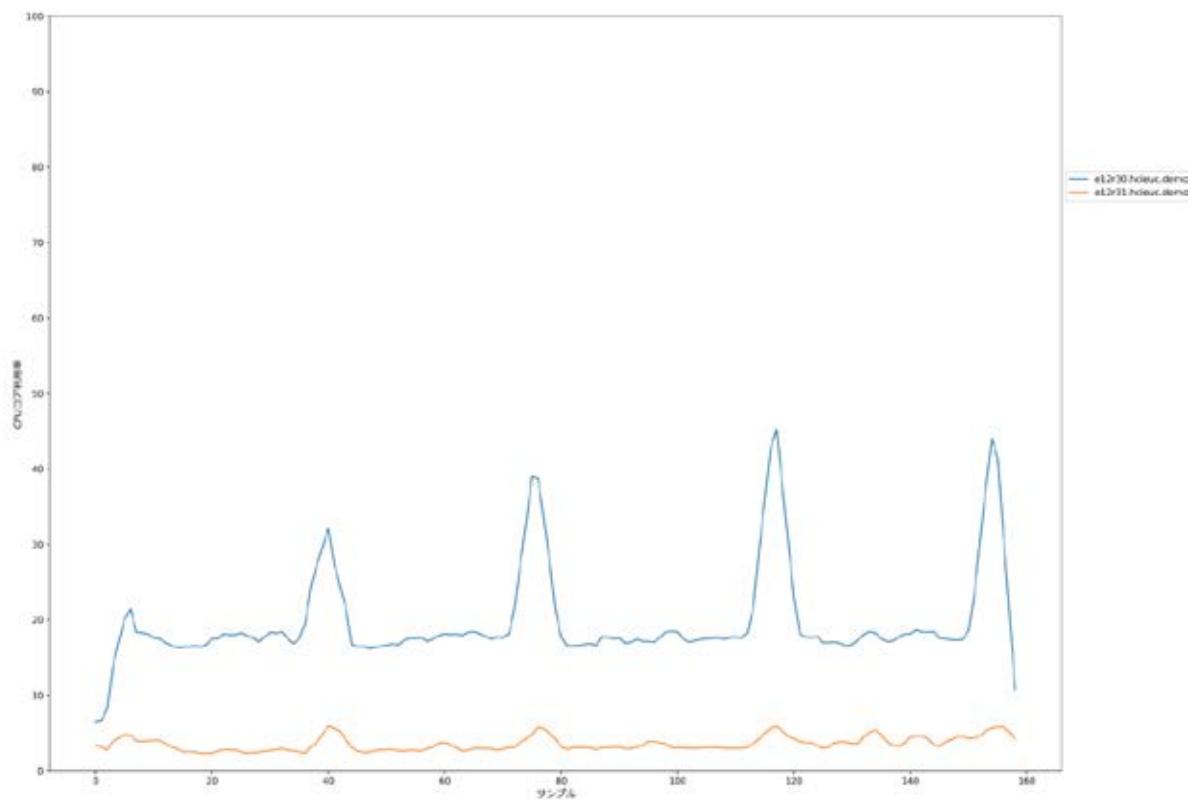


図 51) Showcase : GPU 利用率 – 16Qx1

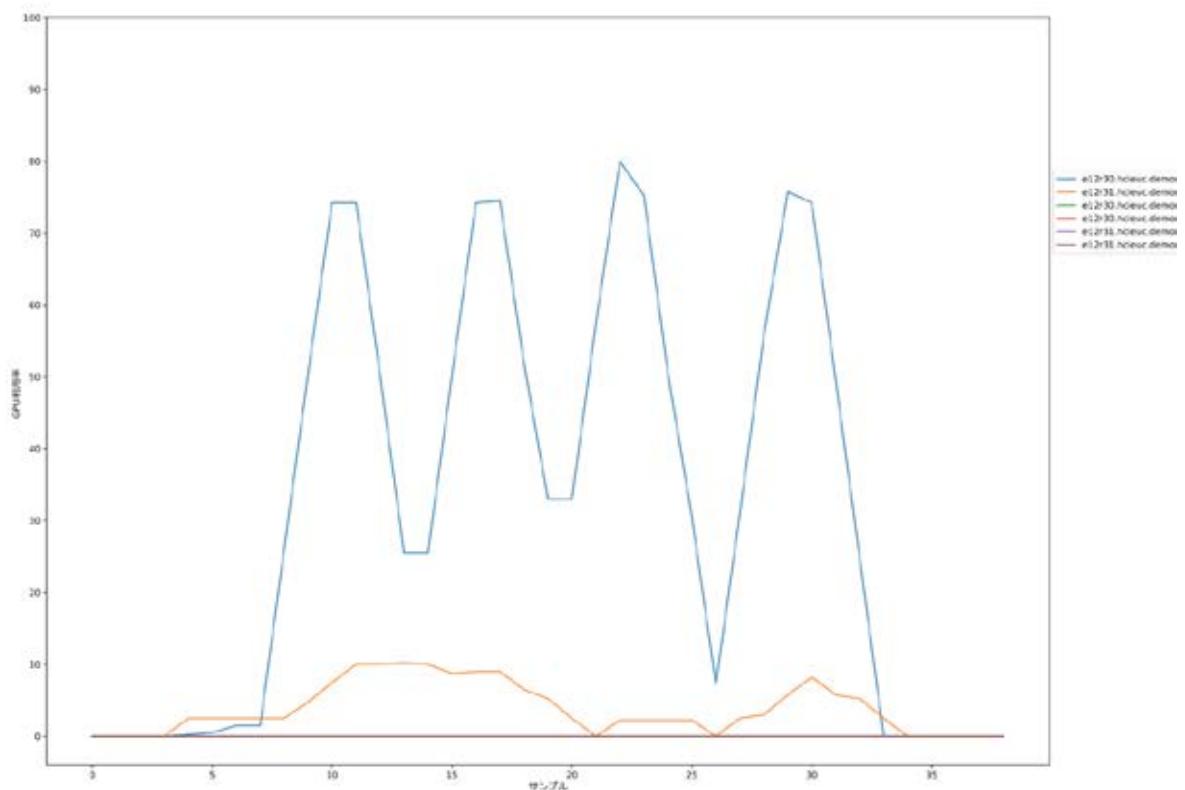


図 52) Showcase : GPU 利用率 – 4Qx12

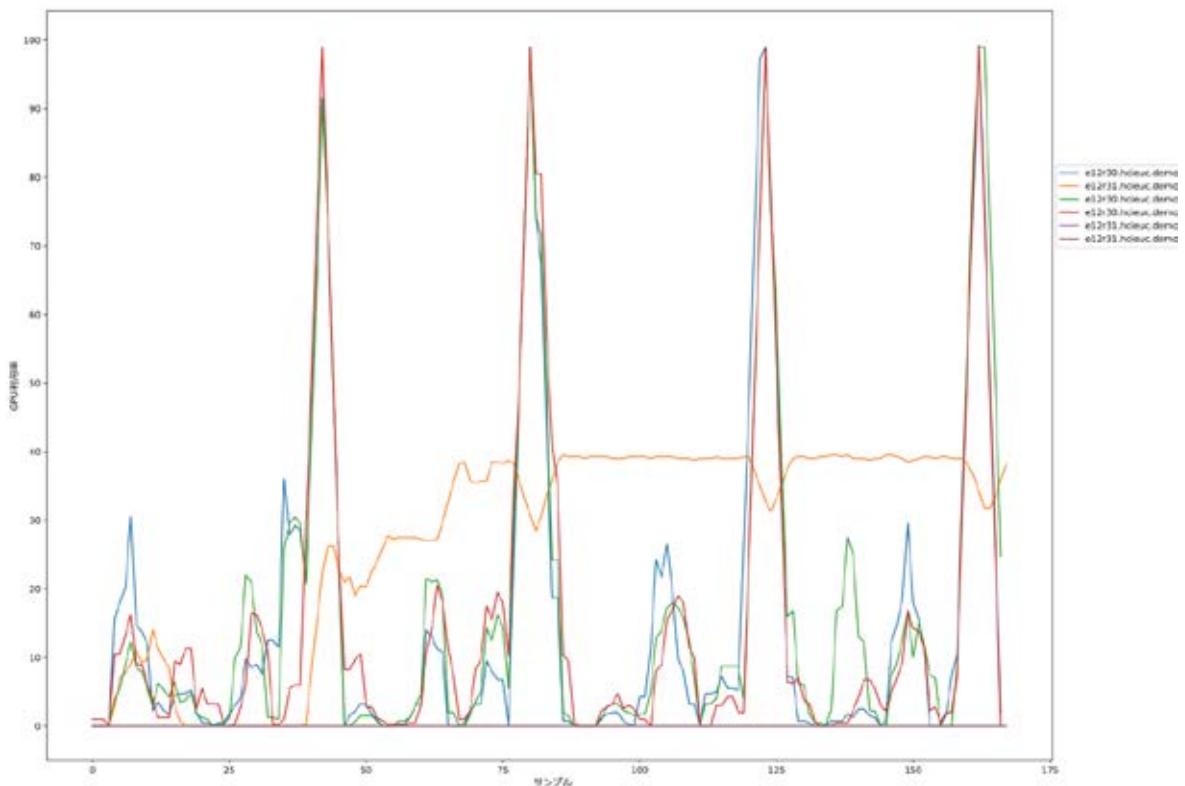


図 53) Showcase のサンプル



6.9 Siemens NX (snx-03)

snx-03 ビューセットは、Siemens PLM の NX 8.0 アプリケーションで生成されたグラフィックスワークフローのトレースから作成されました。モデル サイズは頂点数 715 万～845 万です。

このビューセットには、ワイヤフレーム、アンチエイリアス、シェーディング、エッジ付きのシェーディング、スタジオ モードなど、アプリケーションでサポートされる多数のレンダリング モードが含まれています。

このビューセットには次のテストが含まれます。

このビューセットには次のテストが含まれます。

- パワートレイン (アドバンスト スタジオ モード)
- パワートレイン (シェーディング モード)
- パワートレイン (エッジ付きのシェーディング モード)
- パワートレイン (スタジオ モード)
- パワートレイン (ワイヤフレーム モード)
- SUV (アドバンスト スタジオ モード)
- SUV (シェーディング モード)
- SUV (エッジ付きのシェーディング モード)
- SUV (スタジオ モード)
- SUV (ワイヤフレーム モード)

図 54) Siemens NX の総合スコア

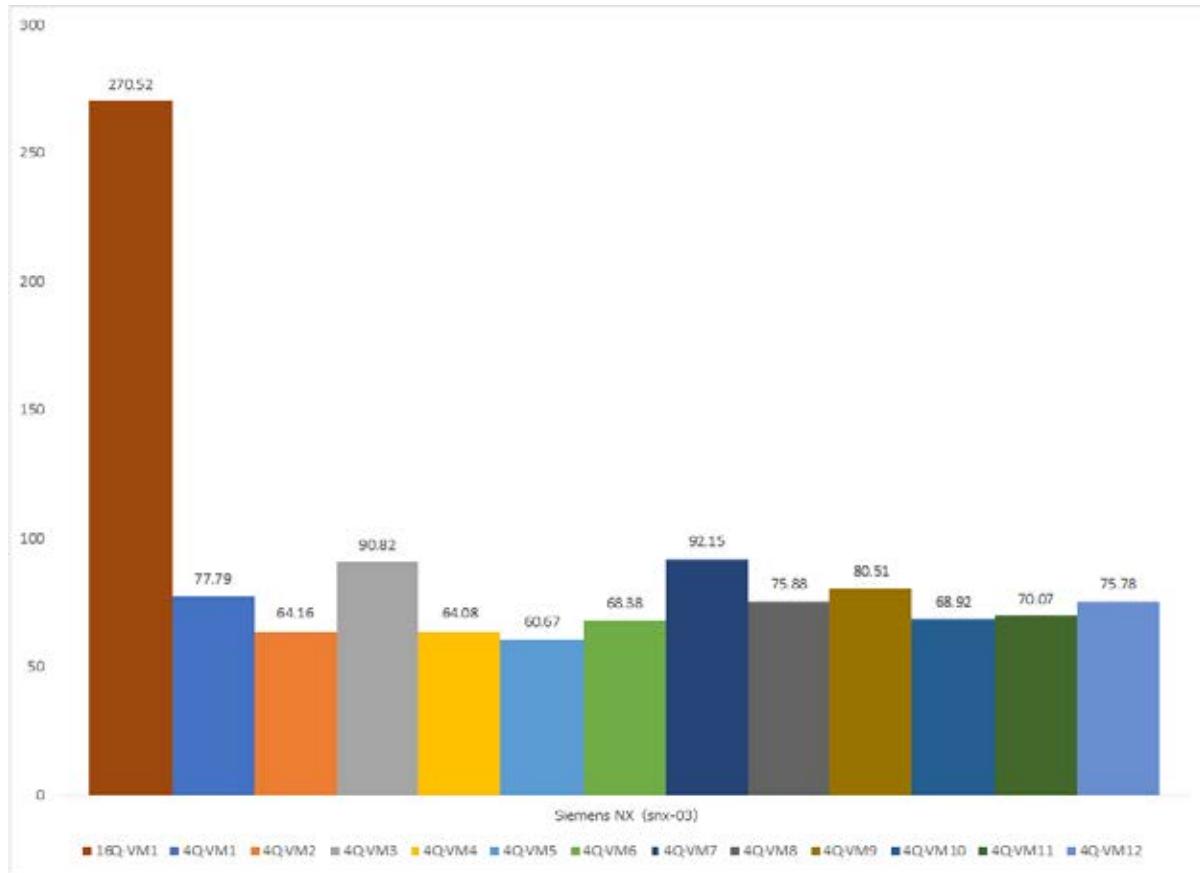


図 55) Siemens NX : vSphere の CPU 利用率 – 16Qx1

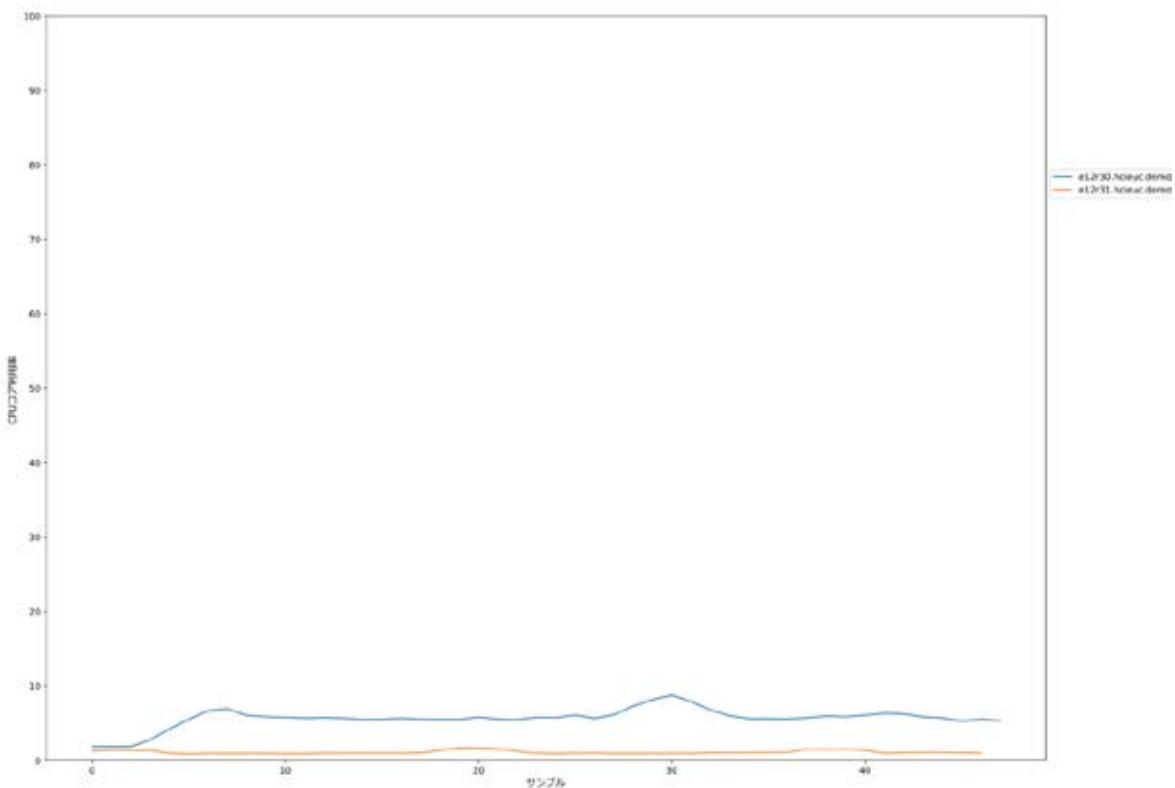


図 56) Siemens NX : vSphere の CPU 利用率 – 4Qx12

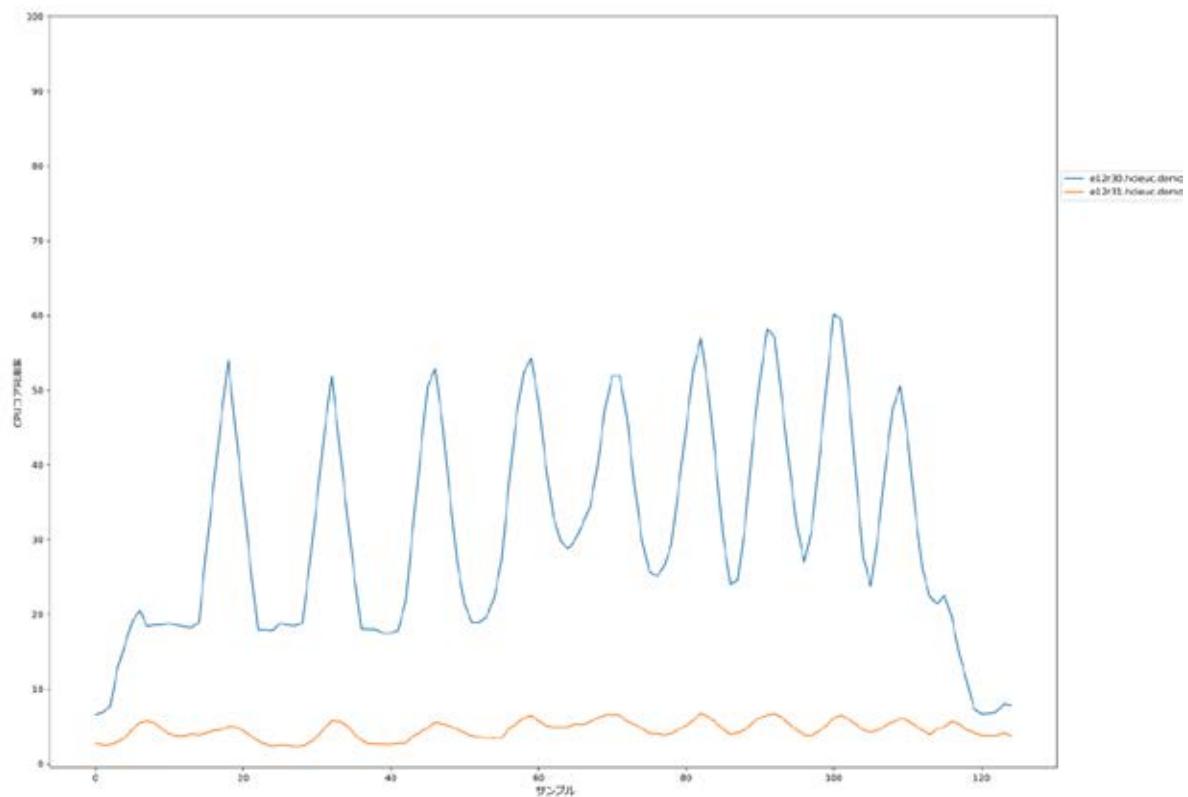


図 57) Siemens NX : GPU 利用率 – 16Qx1

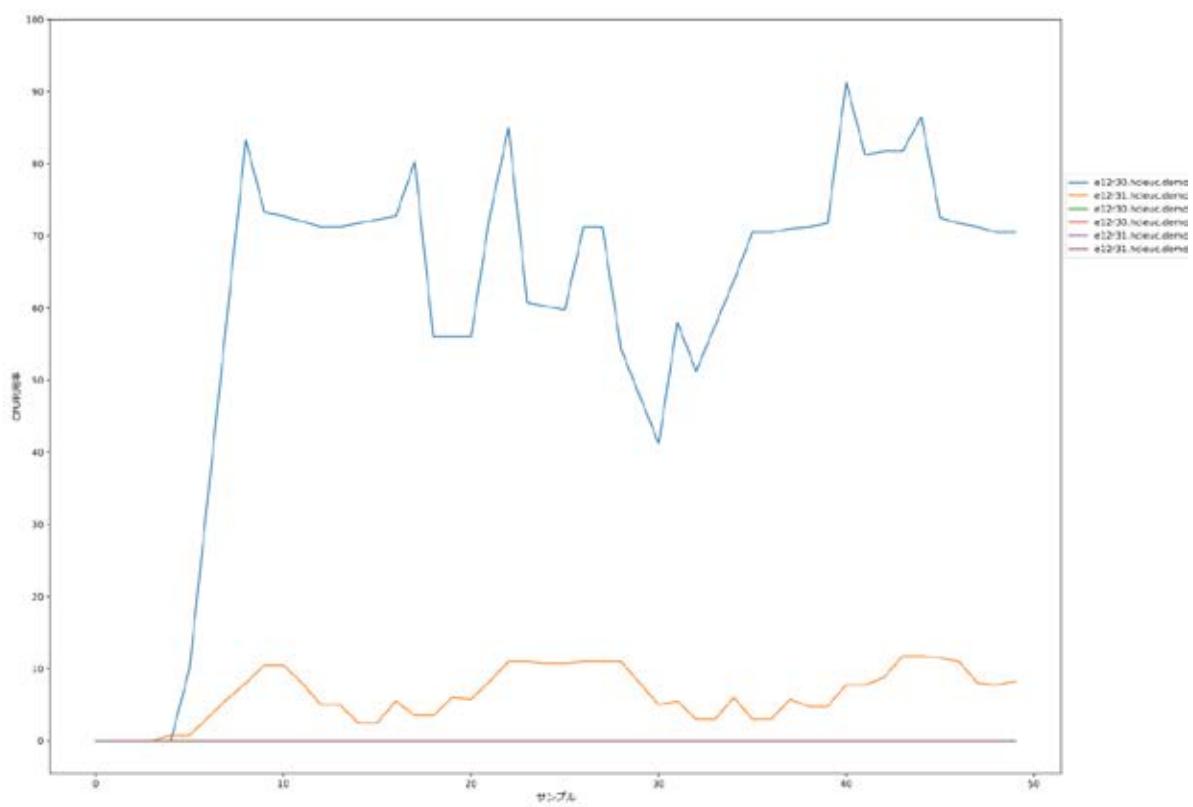


図 58) Siemens NX : GPU 利用率 – 4Qx12

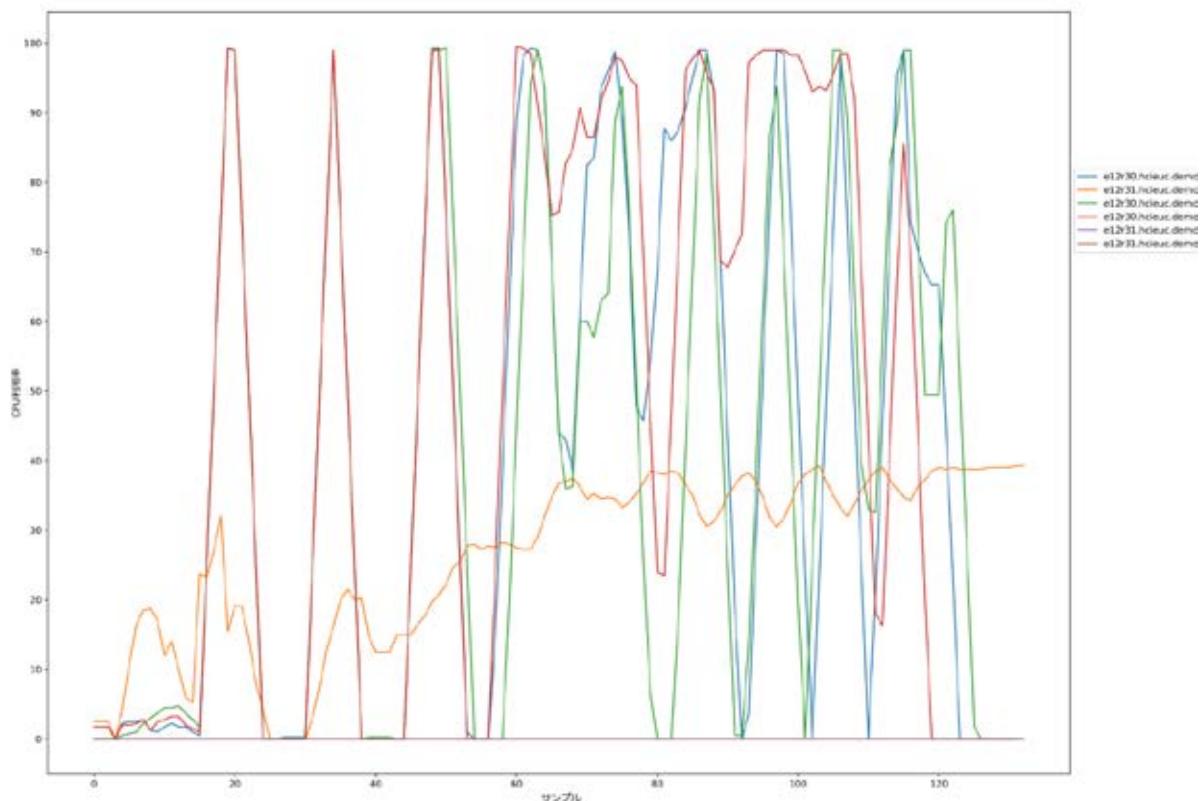
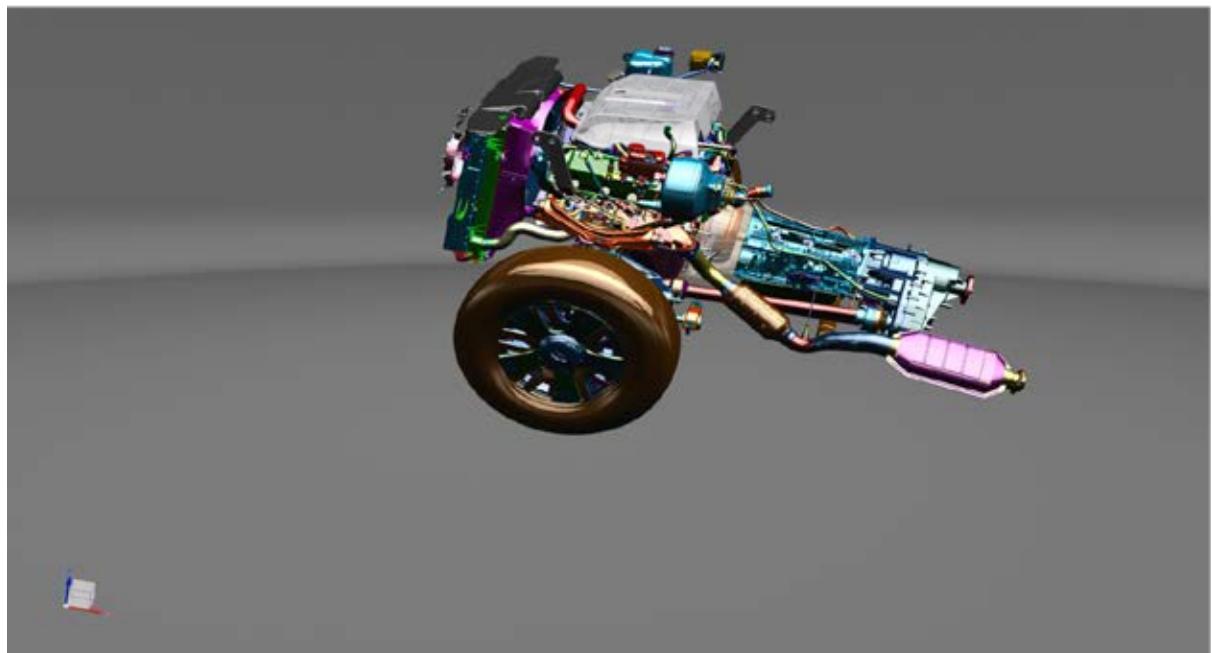


図 59) Siemens NX のサンプル



6.10 SOLIDWORKS (sw-04)

sw-04 ビューセットは、Dassault Systemes の SOLIDWORKS 2013 SP1 アプリケーションのトレースから作成されました。このビューセットで使用されているモデルのサイズは頂点数 210 万～2,100 万です。

このビューセットには、シェーディング モード、エッジ付きシェーディング、アンビエント オクルージョン、シェーダ、環境マップなど、アプリケーションでサポートされる多数のレンダリング モードが含まれています。

このビューセットには次のテストが含まれます。

- 自動車（シェーディング モード）-- 環境キューブマップを使用した通常シェーダ
- 自動車（シェーディング モード）-- 環境キューブマップを使用したバンプ視差マッピング
- 自動車（シェーディング モード）-- 通常シェーダと環境マップでアンビエント オクルージョンを有効化
- 自動車（エッジ付きのシェーディング モード）-- 環境キューブマップを使用した通常シェーダ
- 自動車（ワイヤフレーム モード）
- ラリー カー（シェーディング モード）-- 通常シェーダと環境マップでアンビエント オクルージョンを有効化
- ラリー カー（シェーディング モード）-- 環境キューブマップを使用した通常シェーダ
- ラリー カー（エッジ付きシェーディング モード）-- 環境キューブマップを使用した通常シェーダ
- テスラ タワー（シェーディング モード）-- 通常シェーダと環境マップでアンビエント オクルージョンを有効化
- テスラ タワー（シェーディング モード）-- 環境キューブマップを使用した通常シェーダ
- テスラ タワー（エッジ付きシェーディング モード）-- 環境キューブマップを使用した通常シェーダ

図 60) SOLIDWORKS の総合スコア

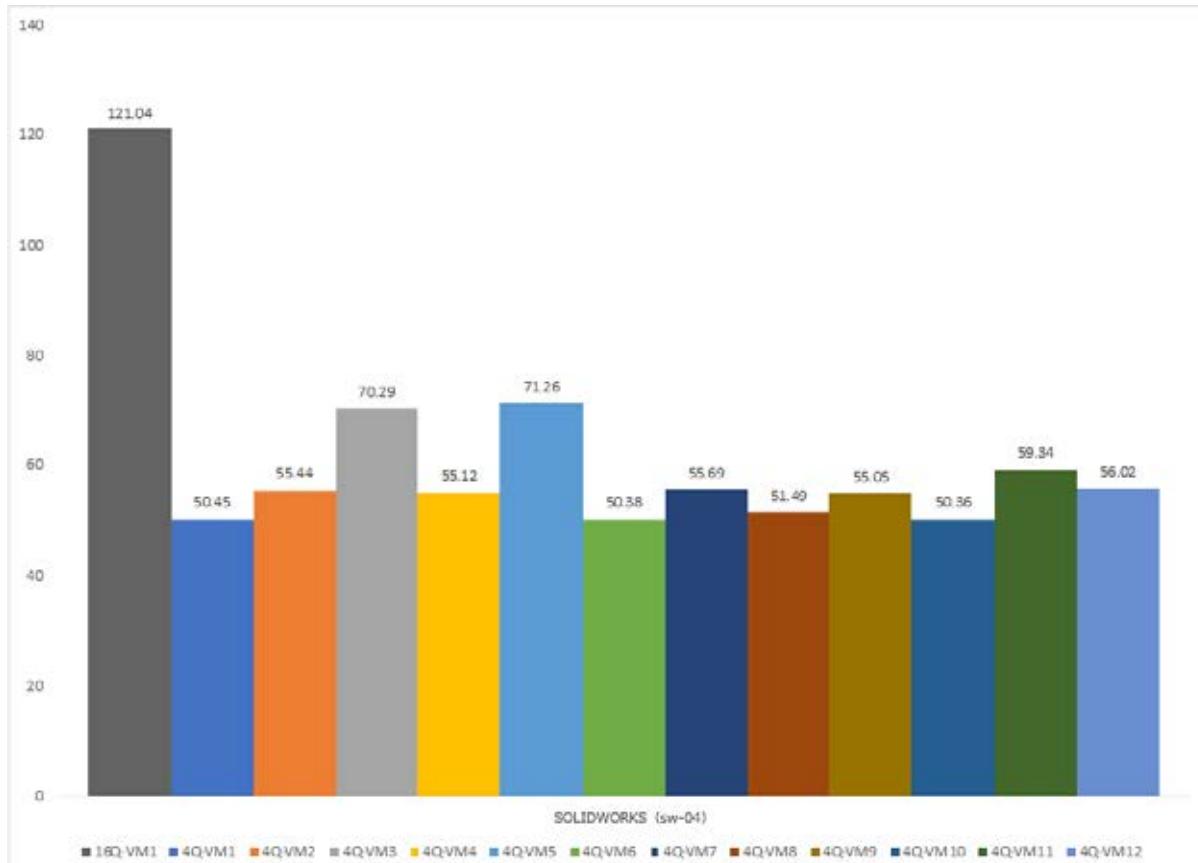


図 61) SOLIDWORKS : vSphere の CPU 利用率 – 16Qx1

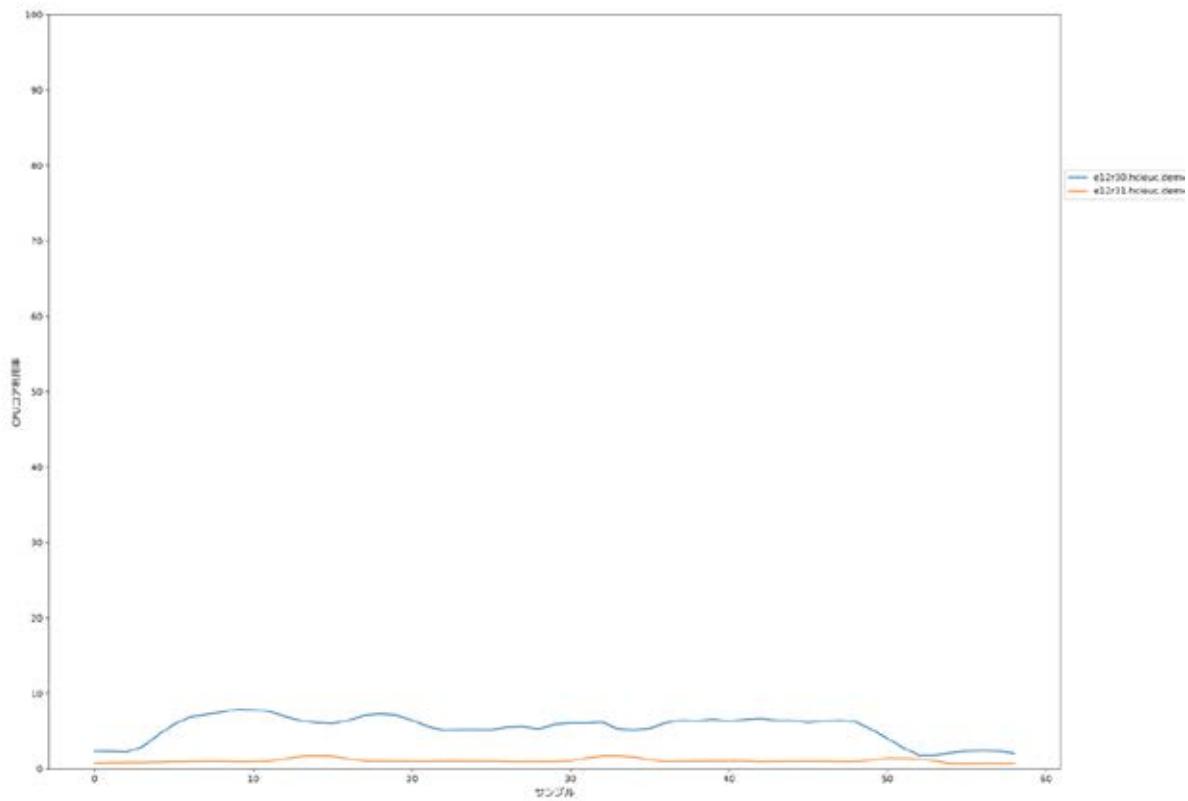


図 62) SOLIDWORKS : vSphere の CPU 利用率 – 4Qx12

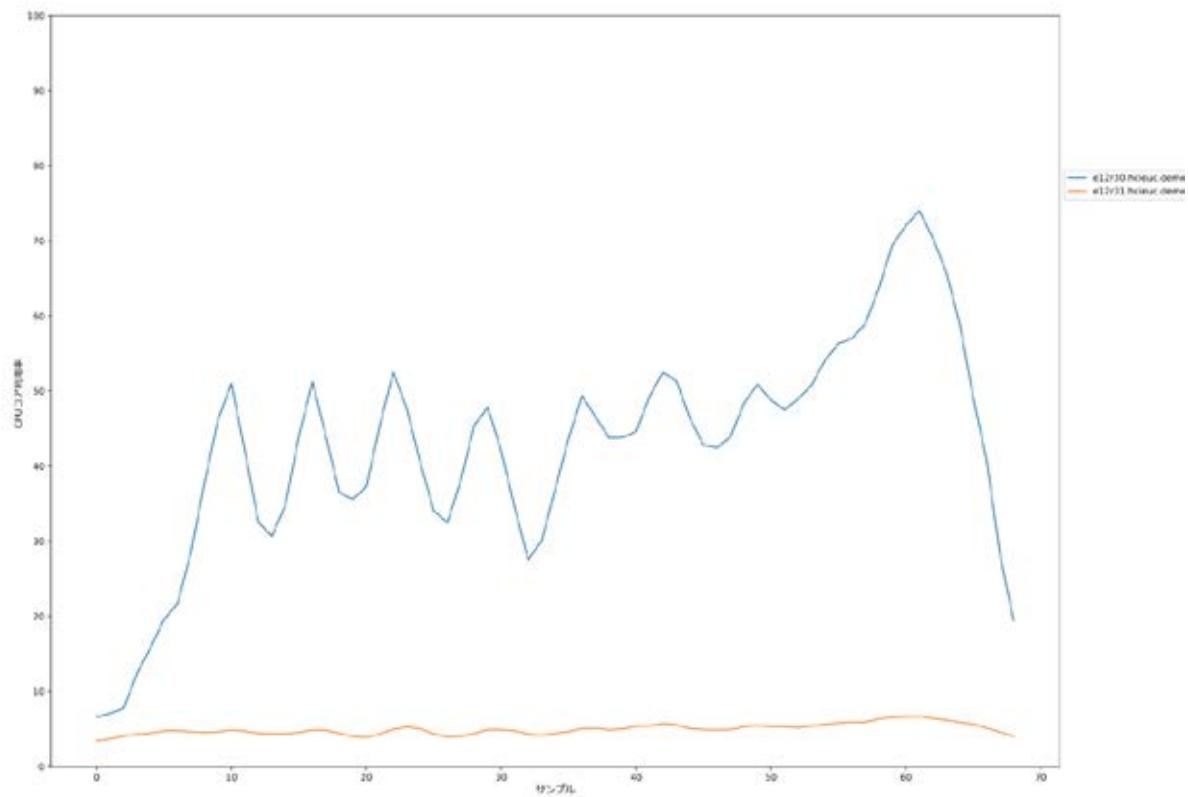


図 63) SOLIDWORKS : GPU 利用率 – 16Qx1

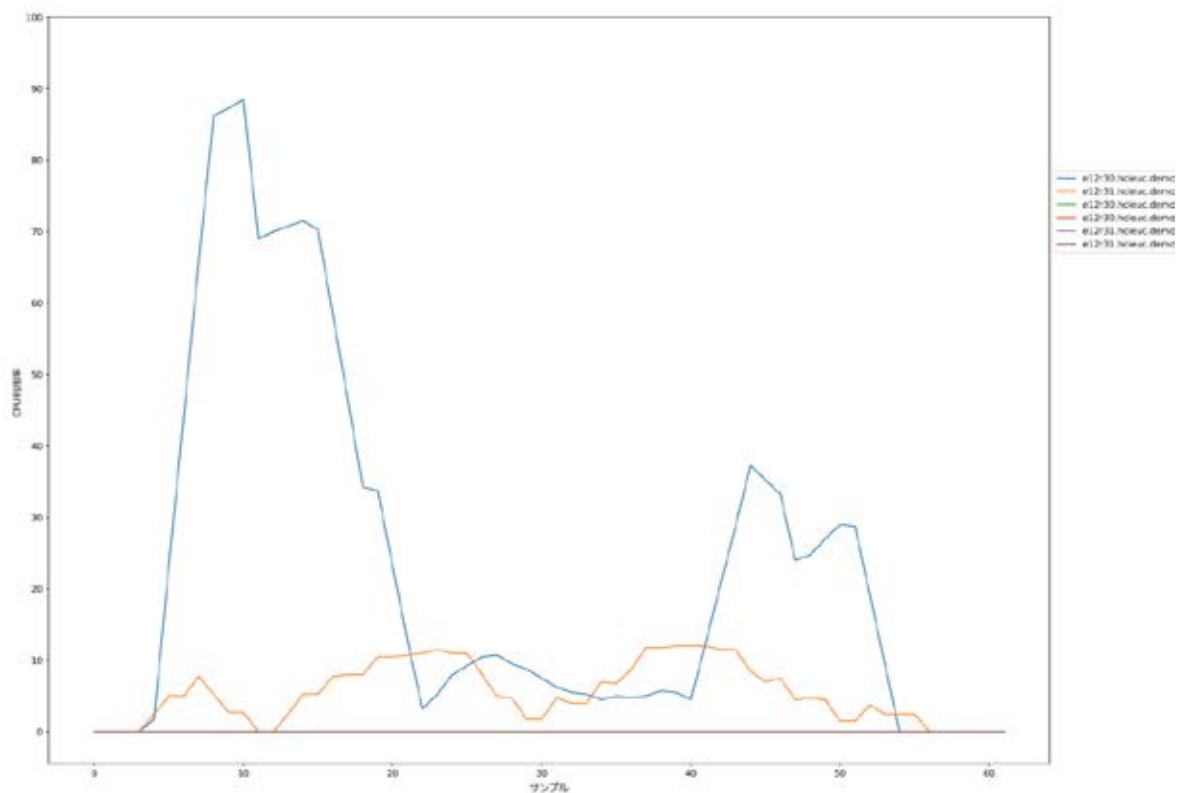


図 64) SOLIDWORKS : GPU 利用率 – 4Qx12

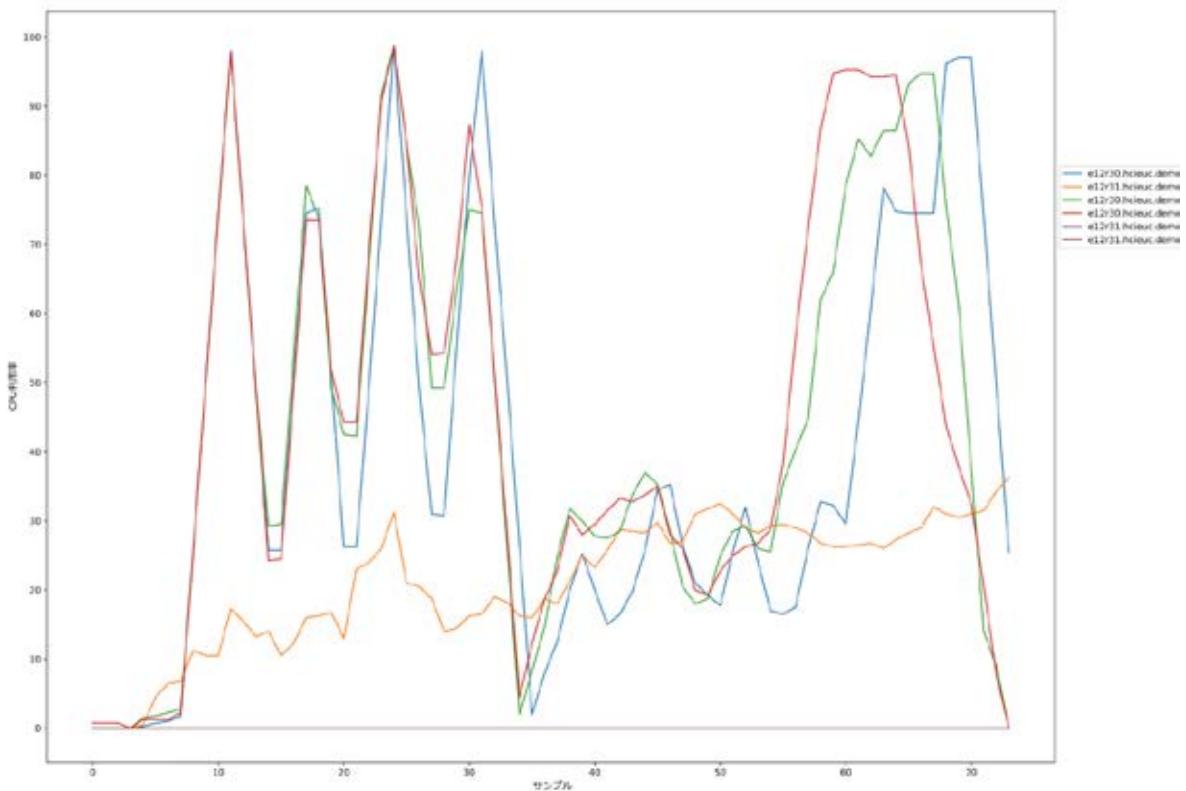


図 65) SOLIDWORKS のサンプル



7 ネットアップ プライベート クラウドとの統合

以下の機能を検討されているお客様には、ネットアップのプライベート クラウドをお勧めします。

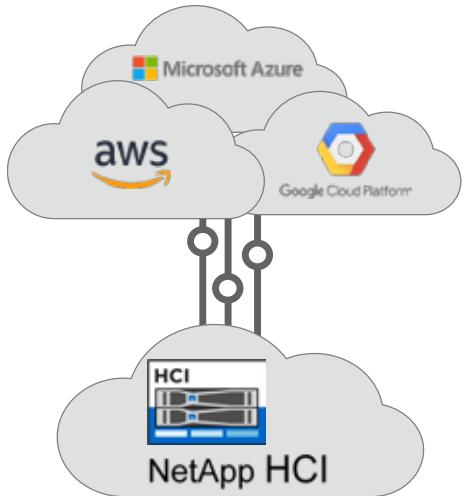
- セルフサービス ポータル
- ポリシー管理による配置やアクセス制御など
- 拡張性に優れたフレームワークによる既存コンポーネントの統合
- 承認ワークフロー
- マシンのライフサイクル管理

ネットアップのプライベート クラウドをすでに導入済みのお客様は、詳細プランを使用してデスクトップ プールを導入し、手動プールとして Horizon にインポートできます。VMware vRealize Orchestrator Plug-in for Horizon 7 がこれらの自動化タスクを支援します。

8 NetApp Kubernetes Service との統合

H615C の GPU リソースをディープ ラーニング推論タスクに使用すれば、GPU リソースを最大限に活用することができます。NVIDIA GPU Cloud はすぐに使用可能なコンテナを備えたコンテナ レジストリであり、このようなタスクの導入時間を短縮します。NetApp Kubernetes Service を一元的な管理プレーンとして使用し、パブリック クラウドまたは NetApp HCI システムに Kubernetes を導入することができます。

図 66) ハイブリッド クラウド



NetApp HCI の管理ノードには Hybrid Cloud Control のポータルが搭載されており、NetApp Kubernetes Service (NKS) もこのポータルから導入できます。

図 67) Hybrid Cloud Control

Cluster Name
HCI-Cluster01
Management Virtual IP
172.21.146.25

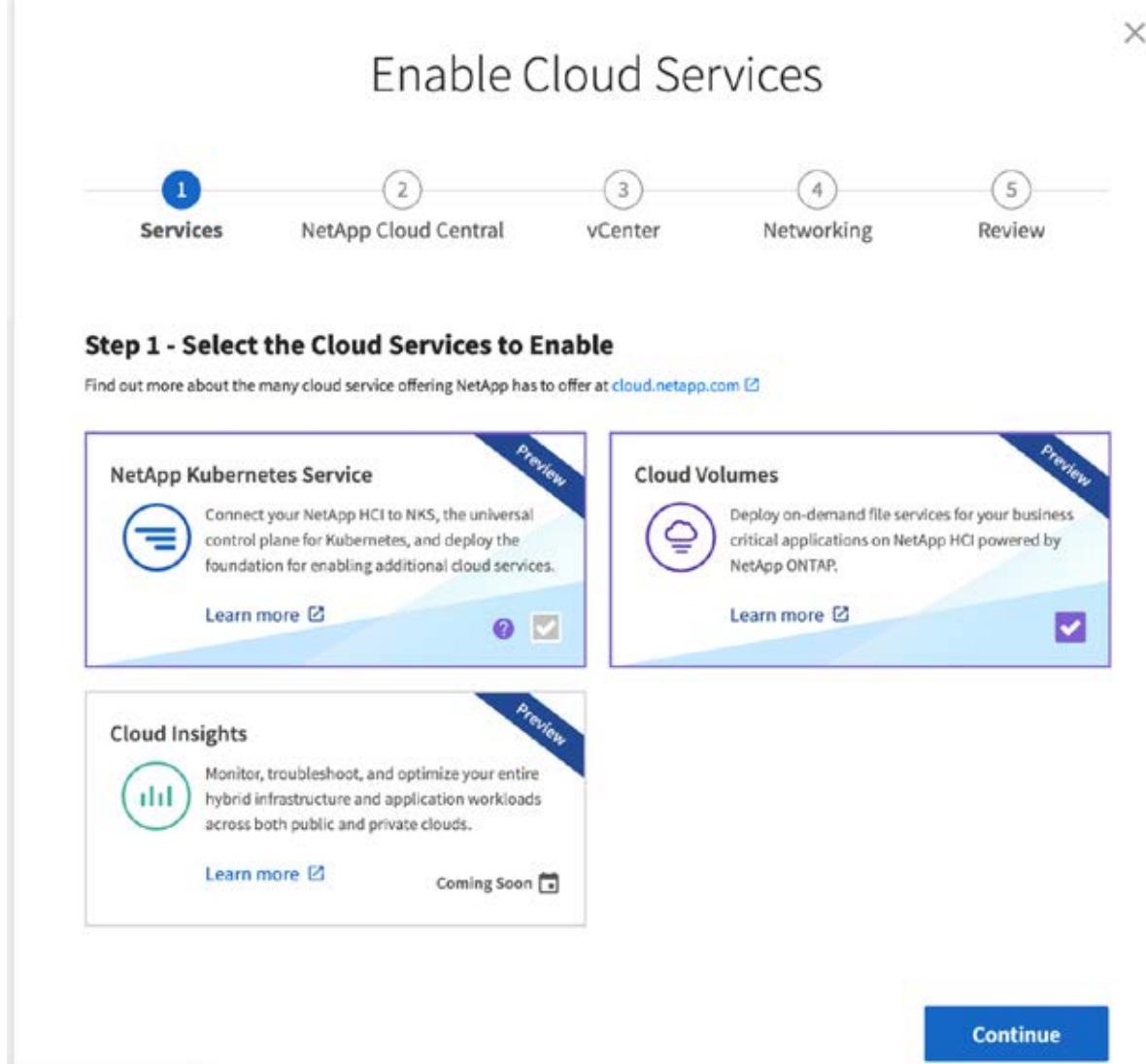
Upgrade Installation
Upgrade your current management node services and the NetApp HCI user interface
Upgrade

Enable NetApp Cloud Services
Register this NetApp HCI installation with Cloud Central and enable NetApp Kubernetes Service and Cloud Volumes on NetApp HCI.
0 of 1 Available Services Enabled
Enable

Expand Installation
Deploy NetApp HCI by adding storage and/or management nodes for your existing installation
Expand

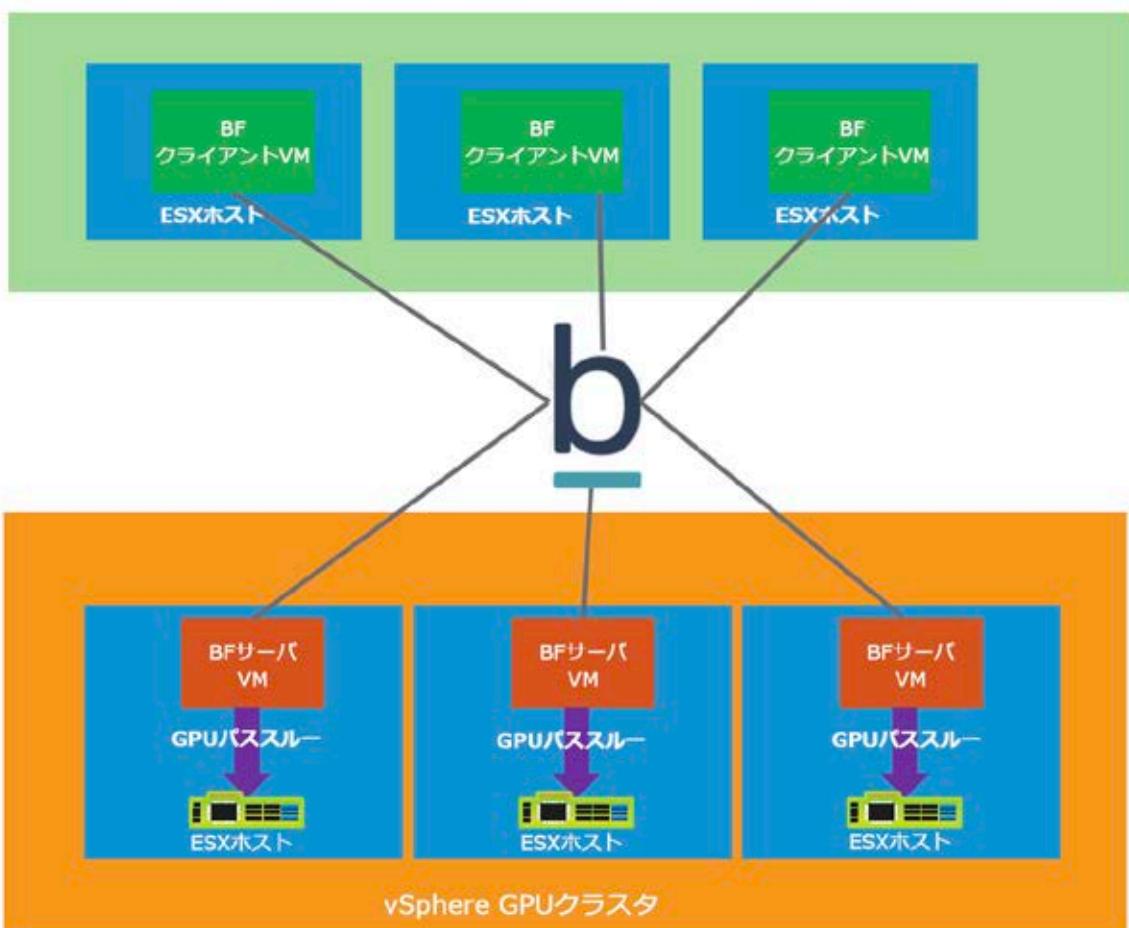
Monitor Installation
Launch Active IQ to view historical and real-time compute and storage metrics
Monitor

図 68) ネットアップ クラウド サービスの有効化



GPU を搭載していないノードでディープ ラーニング推論コンテナを実行する場合は、Bitfusion の FlexDirect を検討してください。FlexDirect はクライアント サーバーアーキテクチャで、サーバーコンポーネントを GPU ノードで実行し、クライアント コンポーネントは IP トранSPORT 経由で通信する別のホストで実行できます。

図 69) Bitfusion FlexDirect



FlexDirect のデモ ビデオは[こちら](#)でご覧いただけます。

Bitfusion は、GPU を搭載していないサーバ CPU に CUDA AI、機械学習、または HPC アプリケーションを導入する場合に最適な選択肢です。ただし、グラフィックスや仮想クライアントコンピューティング (VDI など) はサポートされません。CUDA に完全に対応した AI、機械学習、HPC アプリケーション向けにハイパフォーマンスやワークロード分離が必要な場合は、NVIDIA GPU を搭載したサーバ上でアプリケーションを実行することを推奨します。

9 まとめ

NetApp HCI H615C は、以下を含む幅広いデータセンター ワークロードに最適なシステムです。

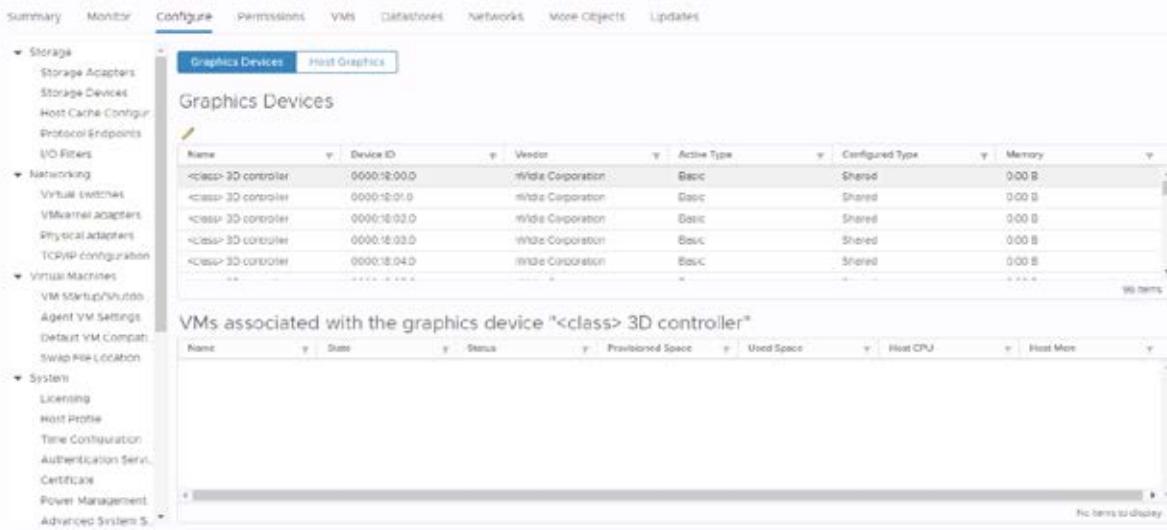
- 最先端のビジネス アプリケーションを使用するナレッジ ワーカー向けの仮想デスクトップ
- 科学者、エンジニア、クリエイティブ プロフェッショナル向けの仮想ワークステーション
- ディープ ラーニングの推論コンピューティング

H615C は、同じラック スペースで H610C よりもユーザ密度が 50% 上昇し、ほとんどのワークロードで 2 倍のパフォーマンスを実現します。

付録 A : vSphere の GPU

vSphere 6.7 Update 1 と T4 GPU を搭載した H615C は、「<class> 3D controller」と呼ばれます。

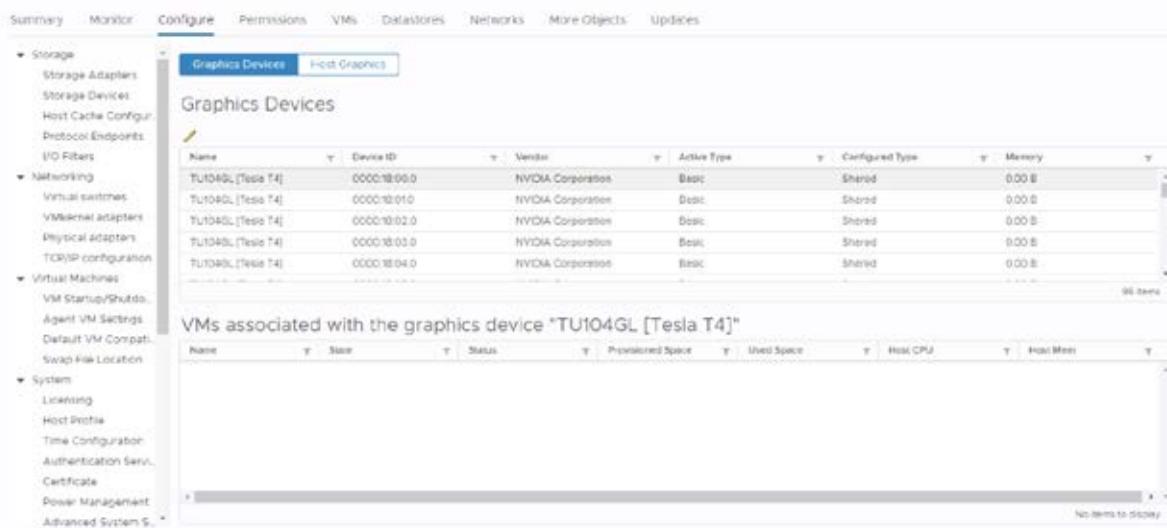
図 70) vSphere 6.7 Update 1 の T4



The screenshot shows the 'Configure' tab selected in the vSphere interface. Under 'Storage', 'Networking', and 'System' categories, there are various adapter and configuration options. The 'Graphics Devices' section is highlighted, showing a table of 'Graphics Devices'. The table lists five entries, all labeled '<class> 3D controller' with Device ID 0000:18:00.0 through 04.0. The 'Active Type' for all entries is 'Basic', and the 'Configured Type' is 'Shared'. The 'Memory' column shows 0.00 B for all entries. Below the table, a section titled 'VMs associated with the graphics device "<class> 3D controller"' is empty, showing 'No items to display'.

H615C は、vSphere 6.7 Update 2 に更新すると、「Tesla T4」と呼ばれます。

図 71) vSphere 6.7 Update 2 の T4



The screenshot shows the 'Configure' tab selected in the vSphere interface. The 'Graphics Devices' section is highlighted, showing a table of 'Graphics Devices'. The table lists five entries, all labeled 'TU104GL [Tesla T4]' with Device ID 0000:18:00.0 through 04.0. The 'Active Type' for all entries is 'Basic', and the 'Configured Type' is 'Shared'. The 'Memory' column shows 0.00 B for all entries. Below the table, a section titled 'VMs associated with the graphics device "TU104GL [Tesla T4]"' is empty, showing 'No items to display'.

NVIDIA vGPU ドライバをインストールしてタイプが[Shared Direct]に変更されたあとも、リブートされるまで[Active Type]は[Shared]のままでです。

図 72) NVIDIA vGPU を使用した T4

The screenshot shows the 'Configure' tab selected in the VMware Horizon 7 interface. Under the 'System' category, the 'Graphics' sub-tab is highlighted. The main content area displays the 'Graphics Devices' table and a list of associated VMs.

Graphics Devices

Name	Device ID	Vendor	Active Type	Configured Type	Memory
NVIDIATesla T4	0000:19:00.0	NVIDIA Corporation	Shared	Shared Direct	14.00 GB
NVIDIATesla T4	0000:19:01.0	NVIDIA Corporation	Basic	Shared Direct	0.00 B
NVIDIATesla T4	0000:19:02.0	NVIDIA Corporation	Basic	Shared Direct	0.00 B
NVIDIATesla T4	0000:19:03.0	NVIDIA Corporation	Basic	Shared Direct	0.00 B
NVIDIATesla T4	0000:19:04.0	NVIDIA Corporation	Basic	Shared Direct	0.00 B

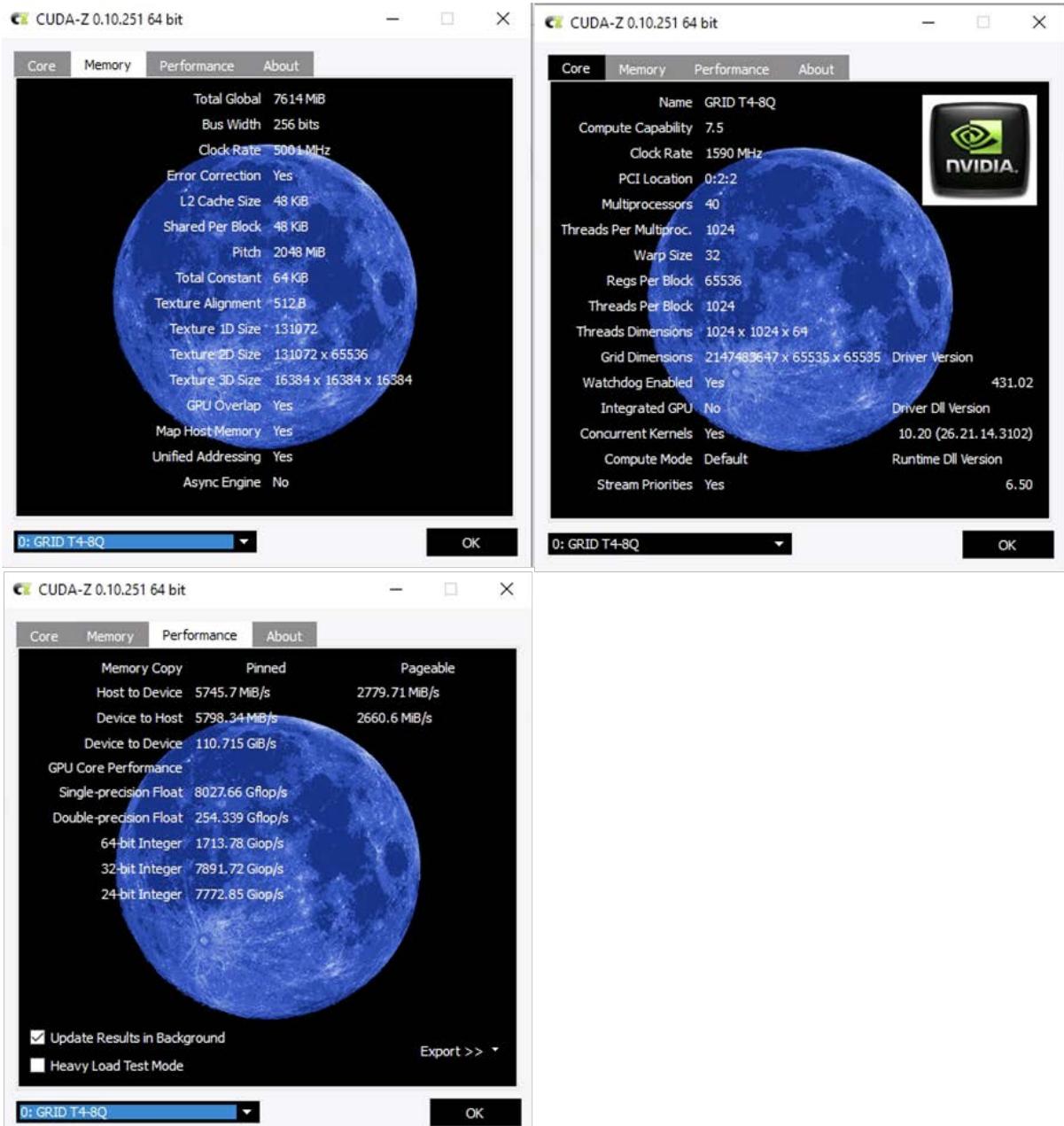
VMs associated with the graphics device "NVIDIATesla T4"

Name	State	Status	Provisioned Space	Used Space	Host CPU	Host Mem
[Empty table row]						

付録 B : T4 CUDA-Z のスクリーンショット

CUDA-Z ツールは CUDA 対応の GPU に関する情報を提供し、<http://cuda-z.sourceforge.net/>からダウンロードできます。図 73 は、vGPU の詳細な仕様を示しています。

図 73) CUDA-Z のスクリーンショット



付録 C : T4 vGPU の設定

ディスプレイ数、特定の GPU プロファイルでサポートされるモニタ解像度、またはフレーム バッファ制限を確認するには、/usr/share/nvidia/vgpu/vgpuconfig.xml ファイルを参照してください。このファイルはあくまで参照用です。直接変更することはしないでください。

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE vgpu SYSTEM "http://www.nvidia.com/dtd/vgpuConfig.dtd">
<vgpuconfig>
  <version>1.0</version>
  <globalSettings>
    <homogeneousVgpus>TRUE</homogeneousVgpus>
    <pluginSoName>libnvidia-vgpu</pluginSoName>
  </globalSettings>
  <vgpuType id="230" name="GRID T4-1Q" class="Quadro">
    <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x130C"/>
    <framebuffer>0x38000000</framebuffer>
    <numHeads>2</numHeads>
    <display width="4096" height="2160"/>
    <mappableVideoSize>0x400000</mappableVideoSize>
    <fbReservation>0x8000000</fbReservation>
    <frlConfig>0x3c</frlConfig>
    <cudaEnabled>0x1</cudaEnabled>
    <eccSupported>1</eccSupported>
    <multiVgpuSupported>0x0</multiVgpuSupported>
    <encoderCapacity>0x64</encoderCapacity>
    <bar1Length>0x100</bar1Length>
    <frame_rate_limiter>1</frame_rate_limiter>
    <license>GRID-Virtual-WS, 2.0;Quadro-Virtual-DWS, 5.0;GRID-Virtual-WS-Ext, 2.0</license>
  </vgpuType>
  <vgpuType id="231" name="GRID T4-2Q" class="Quadro">
    <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x130D"/>
    <framebuffer>0x74000000</framebuffer>
    <numHeads>4</numHeads>
    <display width="4096" height="2160"/>
    <mappableVideoSize>0x400000</mappableVideoSize>
    <fbReservation>0x000000</fbReservation>
    <frlConfig>0x3c</frlConfig>
    <cudaEnabled>0x1</cudaEnabled>
    <eccSupported>1</eccSupported>
    <multiVgpuSupported>0x0</multiVgpuSupported>
    <encoderCapacity>0x64</encoderCapacity>
    <bar1Length>0x100</bar1Length>
    <frame_rate_limiter>1</frame_rate_limiter>
    <license>GRID-Virtual-WS, 2.0;Quadro-Virtual-DWS, 5.0;GRID-Virtual-WS-Ext, 2.0</license>
  </vgpuType>
  <vgpuType id="232" name="GRID T4-4Q" class="Quadro">
    <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x130E"/>
    <framebuffer>0xEC000000</framebuffer>
    <numHeads>4</numHeads>
    <display width="4096" height="2160"/>
    <mappableVideoSize>0x400000</mappableVideoSize>
    <fbReservation>0x14000000</fbReservation>
    <frlConfig>0x3c</frlConfig>
    <cudaEnabled>0x1</cudaEnabled>
    <eccSupported>1</eccSupported>
    <multiVgpuSupported>0x0</multiVgpuSupported>
    <encoderCapacity>0x64</encoderCapacity>
    <bar1Length>0x100</bar1Length>
    <frame_rate_limiter>1</frame_rate_limiter>
    <license>GRID-Virtual-WS, 2.0;Quadro-Virtual-DWS, 5.0;GRID-Virtual-WS-Ext, 2.0</license>
  </vgpuType>
  <vgpuType id="233" name="GRID T4-8Q" class="Quadro">
    <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x130F"/>
    <framebuffer>0x1DC000000</framebuffer>
    <numHeads>4</numHeads>
    <display width="4096" height="2160"/>
    <mappableVideoSize>0x400000</mappableVideoSize>
    <fbReservation>0x24000000</fbReservation>
    <frlConfig>0x3c</frlConfig>
    <cudaEnabled>0x1</cudaEnabled>
    <eccSupported>1</eccSupported>
```

```

<multiVgpuSupported>0x0</multiVgpuSupported>
<encoderCapacity>0x64</encoderCapacity>
<bar1Length>0x100</bar1Length>
<frame_rate_limiter>1</frame_rate_limiter>
<license>GRID-Virtual-WS,2.0;Quadro-Virtual-DWS,5.0;GRID-Virtual-WS-Ext,2.0</license>
</vgpuType>
<vgpuType id="234" name="GRID T4-16Q" class="Quadro">
<devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x1310"/>
<framebuffer>0x3BA400000</framebuffer>
<numHeads>4</numHeads>
<display width="4096" height="2160"/>
<mappableVideoSize>0x400000</mappableVideoSize>
<fbReservation>0x45C00000</fbReservation>
<frlConfig>0x3c</frlConfig>
<cudaEnabled>0x1</cudaEnabled>
<eccSupported>1</eccSupported>
<multiVgpuSupported>0x1</multiVgpuSupported>
<encoderCapacity>0x64</encoderCapacity>
<bar1Length>0x100</bar1Length>
<frame_rate_limiter>1</frame_rate_limiter>
<license>GRID-Virtual-WS,2.0;Quadro-Virtual-DWS,5.0;GRID-Virtual-WS-Ext,2.0</license>
</vgpuType>
<vgpuType id="319" name="GRID T4-4C" class="Compute">
<devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x139A"/>
<framebuffer>0xEC000000</framebuffer>
<numHeads>1</numHeads>
<display width="4096" height="2160"/>
<mappableVideoSize>0x400000</mappableVideoSize>
<fbReservation>0x14000000</fbReservation>
<frlConfig>0x3c</frlConfig>
<cudaEnabled>0x1</cudaEnabled>
<eccSupported>1</eccSupported>
<multiVgpuSupported>0x0</multiVgpuSupported>
<encoderCapacity>0x64</encoderCapacity>
<bar1Length>0x100</bar1Length>
<frame_rate_limiter>1</frame_rate_limiter>
<license>NVIDIA-vComputeServer,9.0;Quadro-Virtual-DWS,5.0</license>
</vgpuType>
<vgpuType id="320" name="GRID T4-8C" class="Compute">
<devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x139B"/>
<framebuffer>0x1DC000000</framebuffer>
<numHeads>1</numHeads>
<display width="4096" height="2160"/>
<mappableVideoSize>0x400000</mappableVideoSize>
<fbReservation>0x24000000</fbReservation>
<frlConfig>0x3c</frlConfig>
<cudaEnabled>0x1</cudaEnabled>
<eccSupported>1</eccSupported>
<multiVgpuSupported>0x0</multiVgpuSupported>
<encoderCapacity>0x64</encoderCapacity>
<bar1Length>0x100</bar1Length>
<frame_rate_limiter>1</frame_rate_limiter>
<license>NVIDIA-vComputeServer,9.0;Quadro-Virtual-DWS,5.0</license>
</vgpuType>
<vgpuType id="321" name="GRID T4-16C" class="Compute">
<devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x1375"/>
<framebuffer>0x3BA400000</framebuffer>
<numHeads>1</numHeads>
<display width="4096" height="2160"/>
<mappableVideoSize>0x400000</mappableVideoSize>
<fbReservation>0x45C00000</fbReservation>
<frlConfig>0x3c</frlConfig>
<cudaEnabled>0x1</cudaEnabled>
<eccSupported>1</eccSupported>
<multiVgpuSupported>0x1</multiVgpuSupported>
<encoderCapacity>0x64</encoderCapacity>
<bar1Length>0x100</bar1Length>
<frame_rate_limiter>1</frame_rate_limiter>
<license>NVIDIA-vComputeServer,9.0;Quadro-Virtual-DWS,5.0</license>
</vgpuType>
<vgpuType id="225" name="GRID T4-1A" class="NVS">
<devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x1311"/>
<framebuffer>0x38000000</framebuffer>
<numHeads>1</numHeads>

```

```

<display width="1280" height="1024"/>
<mappableVideoSize>0x400000</mappableVideoSize>
<fbReservation>0x8000000</fbReservation>
<frlConfig>0x3c</frlConfig>
<cudaEnabled>0x0</cudaEnabled>
<eccSupported>0</eccSupported>
<multiVgpuSupported>0x0</multiVgpuSupported>
<encoderCapacity>0x64</encoderCapacity>
<bar1Length>0x100</bar1Length>
<frame_rate_limiter>1</frame_rate_limiter>
<license>GRID-Virtual-Apps, 3.0</license>
</vgpuType>
<vgpuType id="226" name="GRID T4-2A" class="NVS">
  <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x1312"/>
  <framebuffer>0x74000000</framebuffer>
  <numHeads>1</numHeads>
  <display width="1280" height="1024"/>
  <mappableVideoSize>0x400000</mappableVideoSize>
  <fbReservation>0xC000000</fbReservation>
  <frlConfig>0x3c</frlConfig>
  <cudaEnabled>0x0</cudaEnabled>
  <eccSupported>0</eccSupported>
  <multiVgpuSupported>0x0</multiVgpuSupported>
  <encoderCapacity>0x64</encoderCapacity>
  <bar1Length>0x100</bar1Length>
  <frame_rate_limiter>1</frame_rate_limiter>
  <license>GRID-Virtual-Apps, 3.0</license>
</vgpuType>
<vgpuType id="227" name="GRID T4-4A" class="NVS">
  <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x1313"/>
  <framebuffer>0xEC000000</framebuffer>
  <numHeads>1</numHeads>
  <display width="1280" height="1024"/>
  <mappableVideoSize>0x400000</mappableVideoSize>
  <fbReservation>0x14000000</fbReservation>
  <frlConfig>0x3c</frlConfig>
  <cudaEnabled>0x0</cudaEnabled>
  <eccSupported>0</eccSupported>
  <multiVgpuSupported>0x0</multiVgpuSupported>
  <encoderCapacity>0x64</encoderCapacity>
  <bar1Length>0x100</bar1Length>
  <frame_rate_limiter>1</frame_rate_limiter>
  <license>GRID-Virtual-Apps, 3.0</license>
</vgpuType>
<vgpuType id="228" name="GRID T4-8A" class="NVS">
  <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x1314"/>
  <framebuffer>0x1DC000000</framebuffer>
  <numHeads>1</numHeads>
  <display width="1280" height="1024"/>
  <mappableVideoSize>0x400000</mappableVideoSize>
  <fbReservation>0x24000000</fbReservation>
  <frlConfig>0x3c</frlConfig>
  <cudaEnabled>0x0</cudaEnabled>
  <eccSupported>0</eccSupported>
  <multiVgpuSupported>0x0</multiVgpuSupported>
  <encoderCapacity>0x64</encoderCapacity>
  <bar1Length>0x100</bar1Length>
  <frame_rate_limiter>1</frame_rate_limiter>
  <license>GRID-Virtual-Apps, 3.0</license>
</vgpuType>
<vgpuType id="229" name="GRID T4-16A" class="NVS">
  <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x1315"/>
  <framebuffer>0x3BA400000</framebuffer>
  <numHeads>1</numHeads>
  <display width="1280" height="1024"/>
  <mappableVideoSize>0x400000</mappableVideoSize>
  <fbReservation>0x45C00000</fbReservation>
  <frlConfig>0x3c</frlConfig>
  <cudaEnabled>0x0</cudaEnabled>
  <eccSupported>0</eccSupported>
  <multiVgpuSupported>0x0</multiVgpuSupported>
  <encoderCapacity>0x64</encoderCapacity>
  <bar1Length>0x100</bar1Length>
  <frame_rate_limiter>1</frame_rate_limiter>

```

```

<license>GRID-Virtual-Apps, 3.0</license>
</vgpuType>
<vgpuType id="222" name="GRID T4-1B" class="NVS">
  <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x1309"/>
  <framebuffer>0x38000000</framebuffer>
  <numHeads>4</numHeads>
  <display width="2560" height="1600"/>
  <mappableVideoSize>0x400000</mappableVideoSize>
  <fbReservation>0x8000000</fbReservation>
  <frlConfig>0x2d</frlConfig>
  <cudaEnabled>0x0</cudaEnabled>
  <eccSupported>0</eccSupported>
  <multiVgpuSupported>0x0</multiVgpuSupported>
  <encoderCapacity>0x64</encoderCapacity>
  <bar1Length>0x100</bar1Length>
  <frame_rate_limiter>1</frame_rate_limiter>
  <license>GRID-Virtual-PC, 2.0;GRID-Virtual-WS, 2.0;Quadro-Virtual-DWS, 5.0;GRID-Virtual-WS-Ext, 2.0</license>
</vgpuType>
<vgpuType id="252" name="GRID T4-1B4" class="NVS">
  <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x1345"/>
  <framebuffer>0x38000000</framebuffer>
  <numHeads>1</numHeads>
  <display width="4096" height="2160"/>
  <mappableVideoSize>0x400000</mappableVideoSize>
  <fbReservation>0x8000000</fbReservation>
  <frlConfig>0x2d</frlConfig>
  <cudaEnabled>0x0</cudaEnabled>
  <eccSupported>0</eccSupported>
  <multiVgpuSupported>0x0</multiVgpuSupported>
  <encoderCapacity>0x64</encoderCapacity>
  <bar1Length>0x100</bar1Length>
  <frame_rate_limiter>1</frame_rate_limiter>
  <license>GRID-Virtual-PC, 2.0;GRID-Virtual-WS, 2.0;Quadro-Virtual-DWS, 5.0;GRID-Virtual-WS-Ext, 2.0</license>
</vgpuType>
<vgpuType id="223" name="GRID T4-2B" class="NVS">
  <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x130A"/>
  <framebuffer>0x74000000</framebuffer>
  <numHeads>2</numHeads>
  <display width="4096" height="2160"/>
  <mappableVideoSize>0x400000</mappableVideoSize>
  <fbReservation>0xC000000</fbReservation>
  <frlConfig>0x2d</frlConfig>
  <cudaEnabled>0x0</cudaEnabled>
  <eccSupported>0</eccSupported>
  <multiVgpuSupported>0x0</multiVgpuSupported>
  <encoderCapacity>0x64</encoderCapacity>
  <bar1Length>0x100</bar1Length>
  <frame_rate_limiter>1</frame_rate_limiter>
  <license>GRID-Virtual-PC, 2.0;GRID-Virtual-WS, 2.0;Quadro-Virtual-DWS, 5.0;GRID-Virtual-WS-Ext, 2.0</license>
</vgpuType>
<vgpuType id="224" name="GRID T4-2B4" class="NVS">
  <devId vendorId="0x10de" deviceId="0x1EB8" subsystemVendorId="0x10de" subsystemId="0x130B"/>
  <framebuffer>0x74000000</framebuffer>
  <numHeads>4</numHeads>
  <display width="2560" height="1600"/>
  <mappableVideoSize>0x400000</mappableVideoSize>
  <fbReservation>0xC000000</fbReservation>
  <frlConfig>0x2d</frlConfig>
  <cudaEnabled>0x0</cudaEnabled>
  <eccSupported>0</eccSupported>
  <multiVgpuSupported>0x0</multiVgpuSupported>
  <encoderCapacity>0x64</encoderCapacity>
  <bar1Length>0x100</bar1Length>
  <frame_rate_limiter>1</frame_rate_limiter>
  <license>GRID-Virtual-PC, 2.0;GRID-Virtual-WS, 2.0;Quadro-Virtual-DWS, 5.0;GRID-Virtual-WS-Ext, 2.0</license>
</vgpuType>

```

詳細情報の入手方法

このドキュメントに記載されている情報の詳細については、以下のドキュメントや Web サイトを確認してください。

ネットアップ

- NetApp HCI Documentation Center
<http://docs.netapp.com/hci/index.jsp>
- NetApp HCI ドキュメントリソースページ
<https://www.netapp.com/us/documentation/hci.aspx>
- NetApp HCI Theory of Operations
<https://www.netapp.com/us/media/wp-7261.pdf>
- NVA-1129-DESIGN : VMware End-User Computing with NetApp HCI and NVIDIA GPUs
<https://www.netapp.com/us/media/nva-1129-design.pdf>
- NVA-1129-DEPLOY : NetApp HCI for End-User Computing with VMware and NVIDIA GPUs
<https://www.netapp.com/us/media/nva-1129-deploy.pdf>
- ネットアップの製品ドキュメント
<https://www.netapp.com/jp/documentation/index.aspx>

NVIDIA

- NVIDIA Tesla GPUs for virtualization
<https://www.nvidia.com/content/dam/en-zz/Solutions/design-visualization/solutions/resources/documents1/tesla-gpu-linecard-virtualization-us-nvidia-669786-r7.pdf>
- Virtual Workstation 101
<https://www.nvidia.com/content/dam/en-zz/Solutions/design-visualization/solutions/resources/documents1/Virtual-Workstation-101-Technology-Brief.pdf>
- NVIDIA Virtual GPU Packaging, Pricing and Licensing
<https://images.nvidia.com/content/grid/pdf/Virtual-GPU-Packaging-and-Licensing-Guide.pdf>
- NVIDIA RTX
<https://www.nvidia.com/en-us/design-visualization/technologies/rtx/>
- NVIDIA T4 for Virtualization
https://www.nvidia.com/content/dam/en-zz/Solutions/design-visualization/solutions/resources/documents1/TechBrief_T4.pdf
- NVIDIA vGPU Deployment Guide for VMware Horizon 7.5 on VMware vSphere 6.7
<https://images.nvidia.com/content/pdf/vgpu/guides/vgpu-deployment-guide-horizon-on-vsphere-final.pdf>
- NVIDIA Management and Monitoring
<https://www.nvidia.com/en-us/data-center/virtualization/it-management/>

VMware

- TechZone
<https://techzone.vmware.com/>
- Blast Extreme Display Protocol in VMware Horizon 7
<https://techzone.vmware.com/resource/blast-extreme-display-protocol-vmware-horizon-7>
- Bitfusion FlexDirect のデモ
<https://www.youtube.com/watch?v=BM8XhMXTUQY>

アプリケーション

- Autodesk Maya
<https://www.autodesk.com/products/maya/overview>
- Autodesk 3ds Max
<https://www.autodesk.com/products/3ds-max/overview>
- Autodesk Revit
<https://www.autodesk.com/products/revit/overview>
- Autodesk AutoCAD
<https://www.autodesk.com/products/autocad/overview>
- Dassault Systemes SOLIDWORKS
<https://www.solidworks.com/>
- Dassault Systemes CATIA
<https://www.3ds.com/products-services/catia/>
- PTC Creo
<https://www.ptc.com/en/products/cad/creo>
- Siemens NX
<https://www.plm.automation.siemens.com/global/en/products/nx/>
- Adobe Creative Cloud
<https://www.plm.automation.siemens.com/global/en/products/nx/>

SPEC

- SPECviewperf 13 ベンチマーク
<https://www.spec.org/gwpg/gpc.static/vp13info.html>
- SPEC Consortium
<https://www.spec.org/consortium/>

バージョン履歴

バージョン	日付	ドキュメントバージョン履歴
バージョン 1.0	2019 年 8 月	初版リリース

本ドキュメントに記載されている製品や機能のバージョンがお客様の環境でサポートされるかどうかについては、ネットアップ サポートサイトで [Interoperability Matrix Tool \(IMT\)](#) を参照してください。NetApp IMT には、ネットアップがサポートする構成を構築するために使用できる製品コンポーネントやバージョンが定義されています。サポートの可否は、お客様の実際のインストール環境が公表されている仕様に従っているかどうかによって異なります。

著作権に関する情報

Copyright © 2020 NetApp, Inc. All Rights Reserved. Printed in the U.S. このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を隨時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって保護されている場合があります。

本書に含まれるデータは市販品（FAR 2.101 の定義に基づく）に関係し、データの所有権は NetApp, Inc. にあります。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用権を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc. の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用権については、DFARS 252.227-7015(b)項で定められた権利のみが認められます。

商標に関する情報

NetApp、NetApp のロゴ、および <http://www.netapp.com/jp/legal/netapptmlist.aspx> に記載されているマークは、NetApp, Inc. の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。

TR-4792-0819-JP