

NVIDIA QUADRO RTX

QUADRO



NVIDIA TURING GPU

Turing SM

Up to 16 TFLOPS + 16 TIPS
Concurrent FP & INT Execution
Unified L1 Cache
Variable Rate Shading

Tensor Cores

Up to 130 TFLOPS FP16
Up to 260 TOPS INT8
Up to 500 TOPS INT4

Memory

6MB L2 Cache
Up to 384-bit GDDR 6 @ 14Gbps
Up to 672 GB/sec

RT Cores

Up to 10 Giga Rays/sec
Ray Triangle Intersection
BVH Traversal

NVLink

Up to 100 GB/sec
GPU-GPU Memory Access

Display

Native HDR
8K DisplayPort
Virtual Link

Video

HEVC 8K Real Time Encode
25% Improved Bitrate

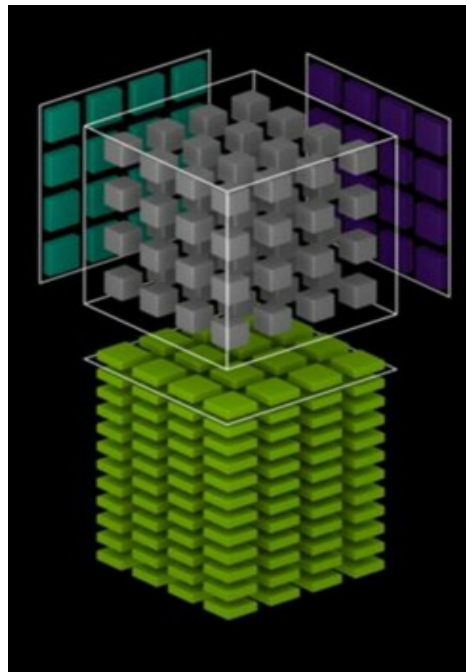
TURING FOR PROFESSIONAL WORKFLOWS

RT CORES



Brings real-time ray tracing to professional graphics workflows

TENSOR CORES

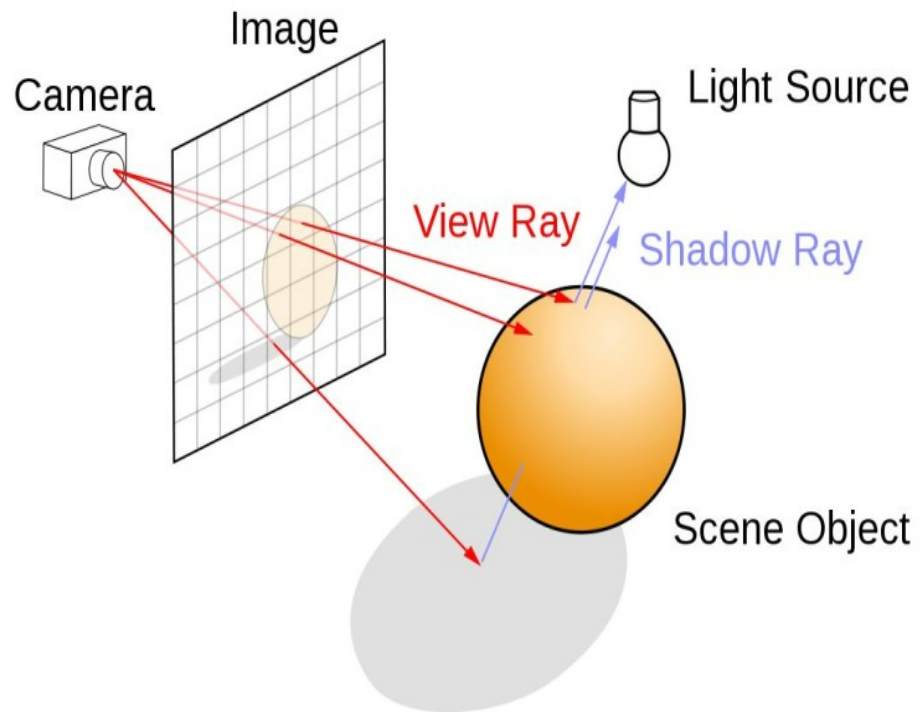


Enables AI-augmented tools and applications

ADVANCED SHADERS

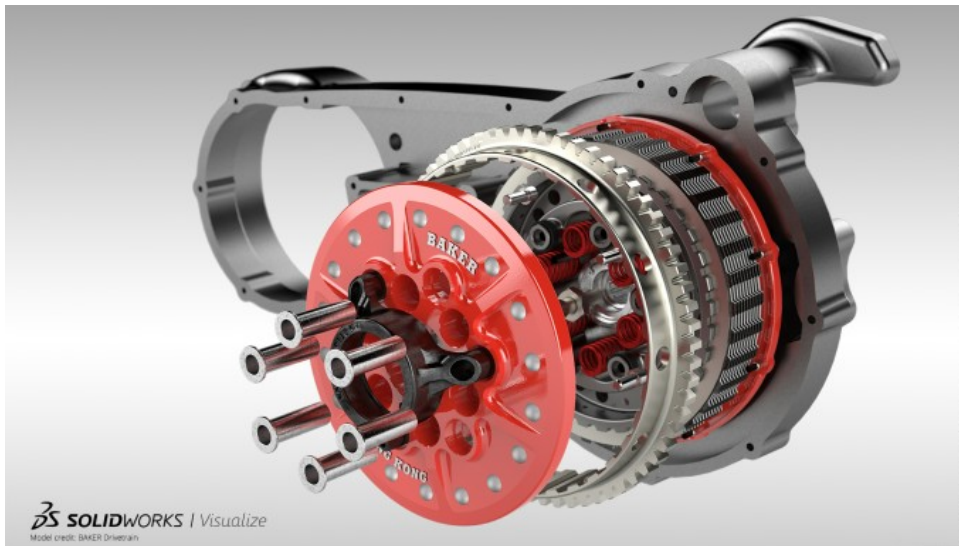
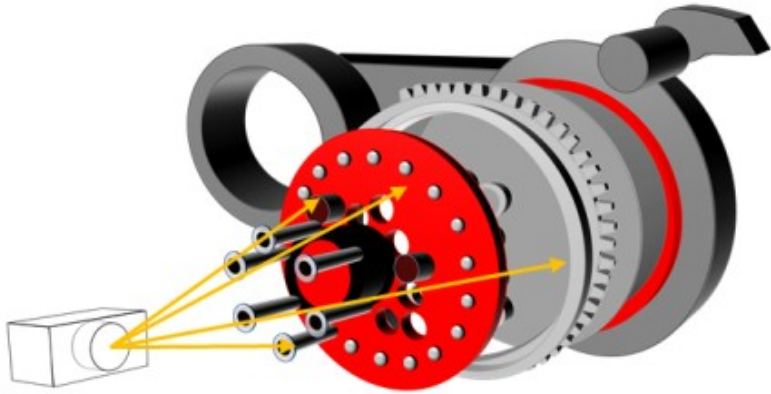
Powers next-generation of graphics, VR, and GPU compute workflows

WHAT IS RAY TRACING?



- Models the behavior of light in the scene
- Produces accurate model of the real world - photorealistic images
- Computationally intensive

TURING RT CORES

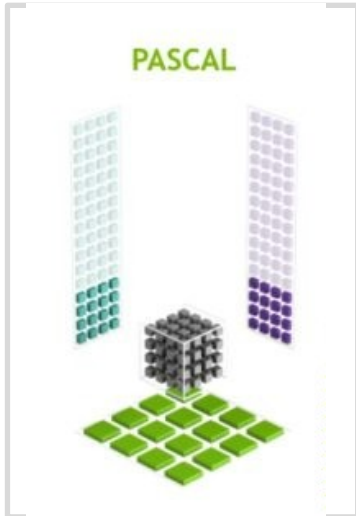


Hardware Accelerated Tracing of Rays Through the Scene

Key Benefits:

- Real-time ray tracing in the application viewport allows for instantaneous feedback and review iteration
- Accelerated offline rendering lets you create photorealistic images faster
- Make better decisions, faster, more iterations without impacting schedules

TURING TENSOR CORES



TURING TENSOR CORE
FP16



TURING TENSOR CORE
INT 8



TURING TENSOR CORE
INT 4



Next Generation of Hardware Accelerated Deep Learning

Key Benefits:

- Turing Tensor Cores deliver fast inferencing performance and support additional precision modes, which boosts inferencing workload performance
- Bring new techniques like Deep learning Super Sampling (DLSS) to your workstation via hardware-accelerated deep learning enabled tools and applications

TURING ADVANCED SHADERS

Mesh Shading

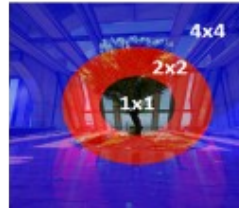
Enables developers to eliminate CPU draw call bottlenecks and use more efficient algorithms to draw triangles.



Mesh Shading to render thousands of objects in real time

Variable Rate Shading (VRS)

More control over pixel shading rate; efficient for effects like motion, blur, foveated shading.



Foveated Shading

Texture Space Shading

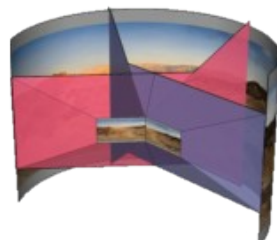
Decouples shading from screen space, improving shading efficiency & reuse.



Shading Re-Use

Multi-View Rendering (MVR)

Extension of Single Pass Stereo rendering multiple views in a single pass with unique view origin positions or view directions.



200° FOV HMD using MVR

Advanced Graphics Technology

Key Benefits:

- Create more objects per scene with more flexible control over the level of detail
- Finer control over shading allows for more dynamic geometry manipulation, letting developers deploy new, optimized algorithms
- Enhancements to single-pass stereo provide greater flexibility and support for new generation of HMDs

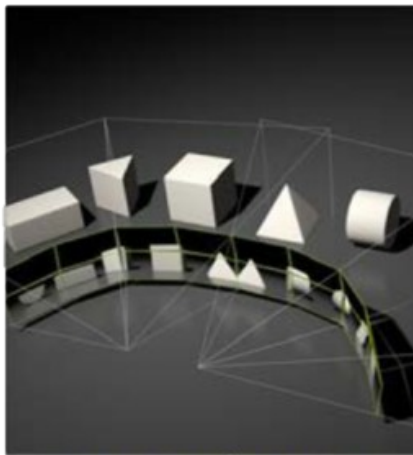
TURING VR



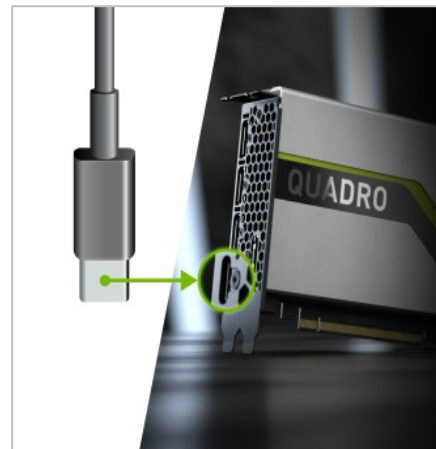
VISUAL QUALITY



ULTRA WIDE FIELD OF VIEW



ACOUSTIC SIMULATION



EASY SETUP

Turing GPU features enhance VR

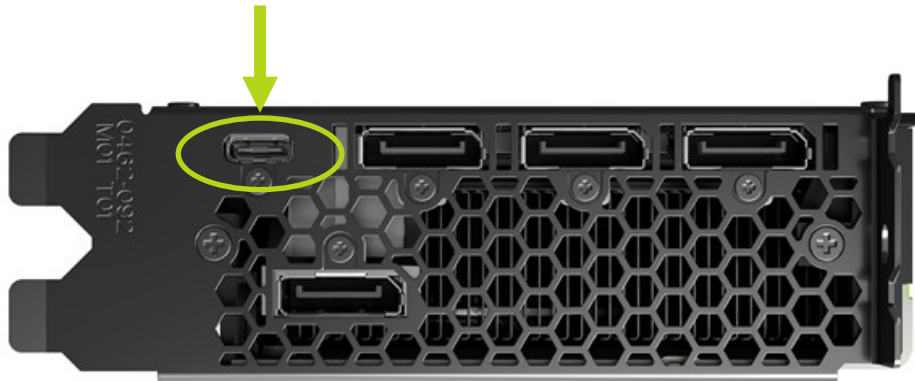
Key Benefits:

- Optimize resolutions with variable rate shading and foveated rendering
- Multi-view rendering provides a wider field of view and support for next-gen HMDs & displays
- RT Cores enable accurate acoustic simulations to deliver more realistic virtual environments
- Easier set up with VirtualLink™ single cable connection

QUADRO RTX VIRTUALLINK™ *

VirtualLink

VirtualLink USB-C Port



VirtualLink is an industry standard Alternate Mode of USB Type-C™ designed to deliver the power, display, and data required to power VR headsets through a single USB Type-C connector.

- 4 lanes HBR3 DisplayPort
- USB 3.1 Gen2 SuperSpeed
- 27 W power
- Industry consortium includes: NVIDIA, VALVE, Oculus, AMD, Microsoft
- [virtuallink.org](http://www.virtuallink.org) for more details

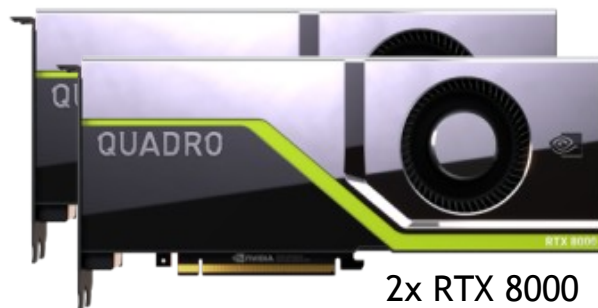
*In preparation for the emerging VirtualLink standard, Turing GPUs have implemented hardware support according to the “VirtualLink Advance Overview”. To learn more about VirtualLink, please see <http://www.virtuallink.org>.

QUADRO RTX NVLINK

High-speed GPU interconnect

Key Benefits:

- Scaled memory and performance lets you split workloads efficiently across two GPUs, sharing up to 96 GB of memory capacity
- Increased bandwidth enables new, advanced SLI display topologies that were previously impossible with PCIe-based solutions



2x RTX 8000
96 GB



2x RTX 6000
48 GB



2x RTX 5000
32 GB

PERFORMANCE

QUADRO NVLINK

Quadro Family NVLink Bridges



Quadro RTX boards only require 1 NVLink bridge

Quadro GPU	NVLink Bridge	Slot Configuration	Bandwidth	Bridges Required
Quadro RTX 8000 RTX 6000	Quadro RTX NVLink HB 2-Slot	2-Slot	Up to 100 GB/s	1
	Quadro RTX NVLink HB 3-Slot	3-Slot		
Quadro RTX 5000	Quadro RTX NVLink 2-Slot	2-Slot	Up to 50 GB/s	1
	Quadro RTX NVLink 3-Slot	3-Slot		
Quadro GV100	NVLink GV100	2-Slot	Up to 200 GB/s	2
Quadro GP100	NVLink GP100	2-Slot	Up to 160 GB/s	2

Bridges are product specific, not cross-compatible

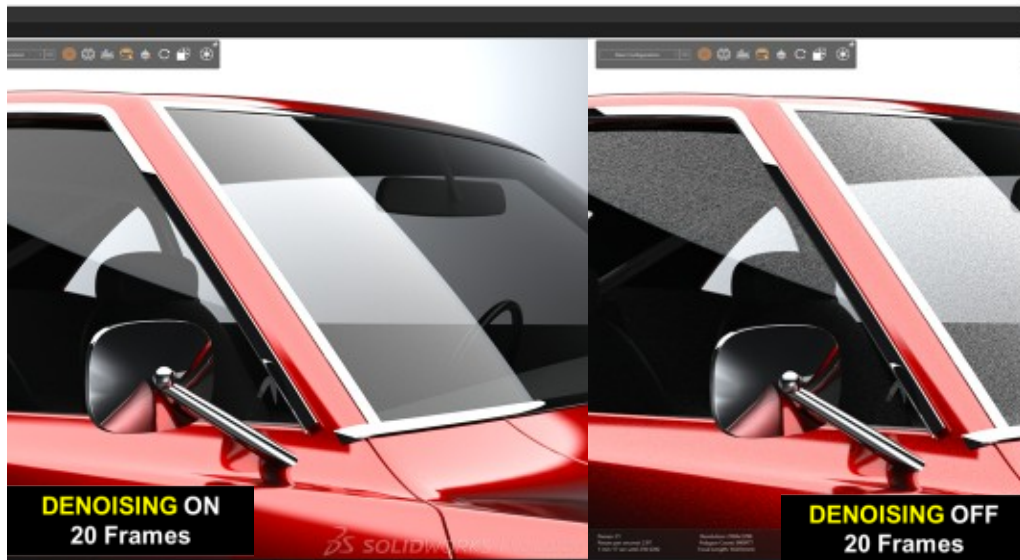


GV100/GP100 boards require 2 NVLink bridges

QUADRO RTX FOR AI

Quadro RTX ideal for AI augmented professional applications and professional AI inferencing deployments

Pro Applications Inferencing

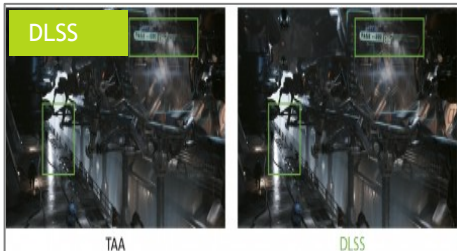
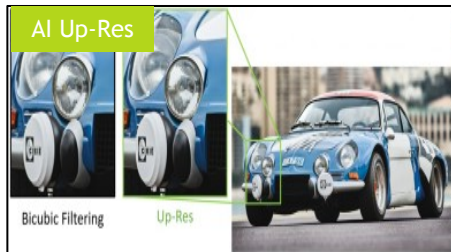


Aggregation Inferencing At-The-Edge



QUADRO RTX FOR AI - NGX

NGX AI-based features




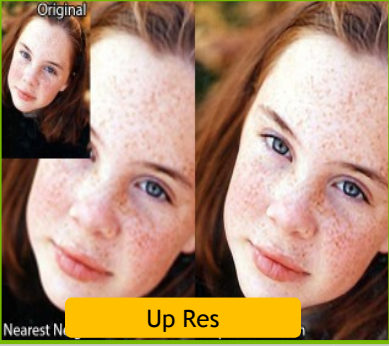
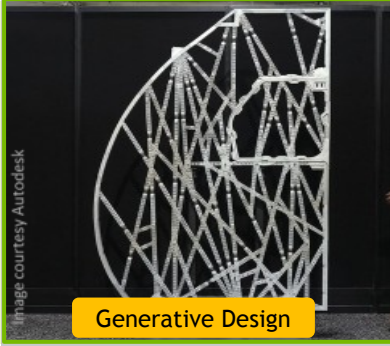
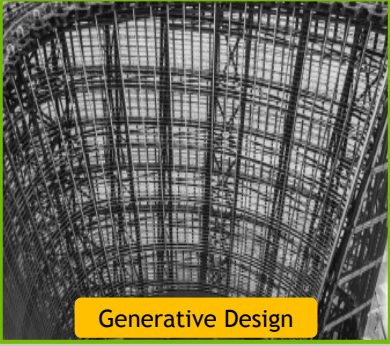





The NVIDIA NGX SDK makes it easy for developers to integrate AI features into their applications with pre-trained neural networks. NGX provide AI-augmented features for video and image processing including:

- **AI InPainting**
Allows the removal of existing content from images and replaces it with realistic computer-generated alternatives.
- **AI Up-Res**
Increases the resolution of an image or video by 2x, 4x or 8x using AI to create new pixels by interpreting the image & intelligently placing data in the new image.
- **DLSS: (Deep Learning Super Sample)**
Removes jagged lines to smooth images, producing a higher quality image faster than by using other techniques.
- **AI Slow-Motion**
Inserts interpolated frames into a video stream to provide smooth, slow-motion video

Details on the NGX SDK: developer.nvidia.com/rtx/ngx

QUADRO RTX VALUE FOR INDUSTRIES

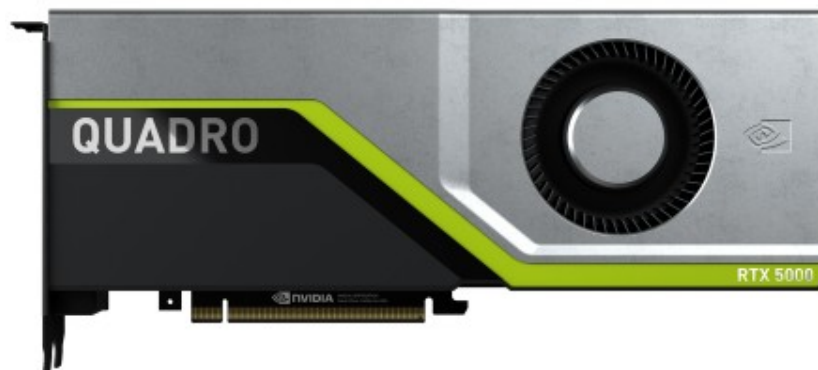
	MEDIA & ENTERTAINMENT	MANUFACTURING	AEC	
RENDERING	 <p>Content Creation</p>	 <p>Product Design</p>	 <p>Building Design</p>	Real-time rendering speeds up the creative workflow
AI	 <p>Up Res</p>	 <p>Generative Design</p>	 <p>Generative Design</p>	AI-augmented tools accelerate the creative process
VR	 <p>Content Creation</p>	 <p>Design Review</p>	 <p>Design Review</p>	VR powers design reviews, compelling content creation and entertainment experiences

QUADRO

QUADRO RTX



RTX 6000



RTX 5000



RTX 4000

QUADRO RTX 6000 KEY SPECIFICATIONS

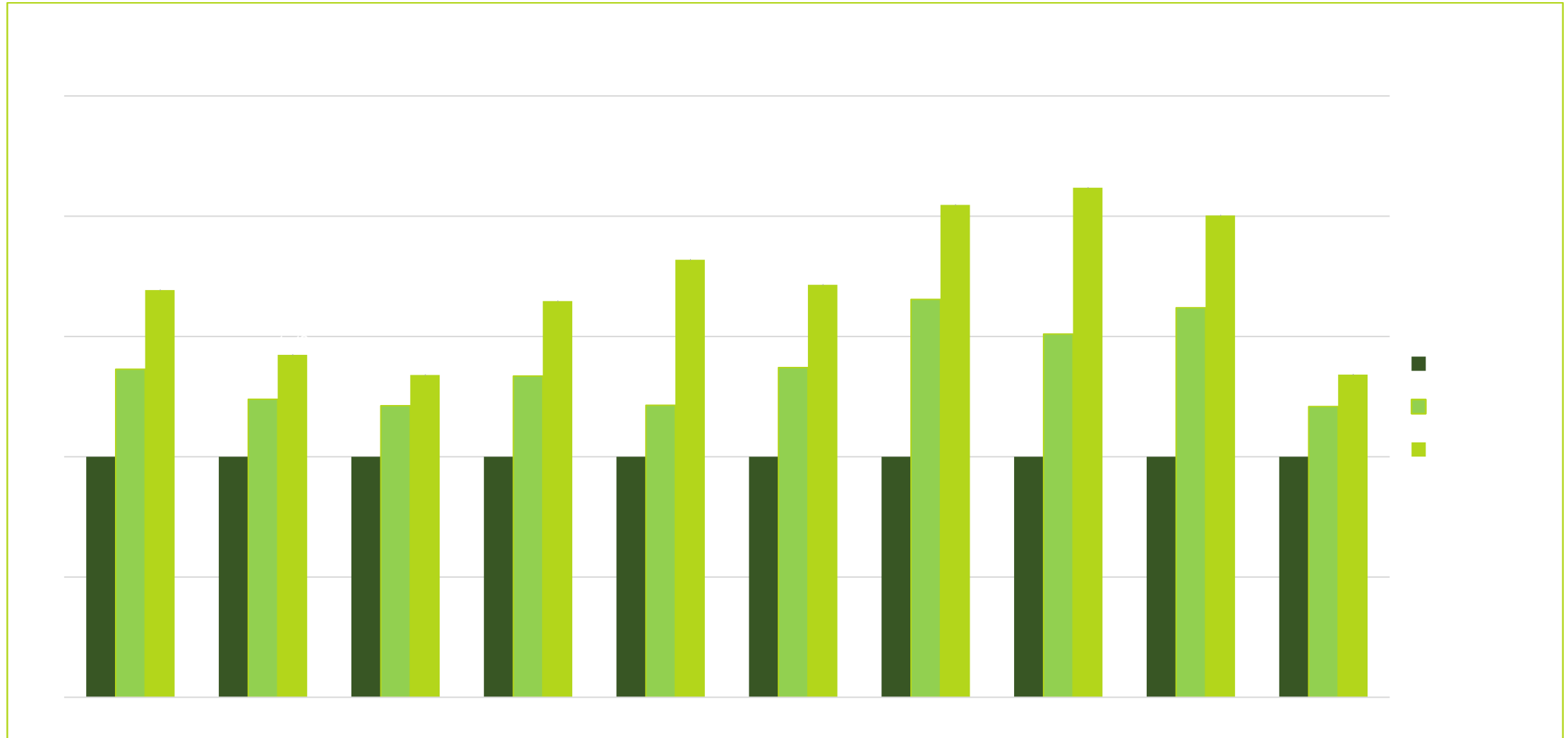


GPU Architecture	Turing
CUDA Cores	4608
RT Cores	72
Tensor Cores	576
Memory Size	24 GB GDDR6
Memory BW	Up to 672 GB/s
NVLink	2-way (2 & 3slot) 100 GB/s bidirectional
Display Support	4x DP + 1x VirtualLink
VR Ready	Yes
VirtualLink™	Yes
Advanced Display	SYNC 2
Board Power	Total Board Power: 295W Total Graphics Power: 260W
Power Connectors	1x 8-pin, 1x 6-pin PCIe

UPGRADING TO RTX 6000

	RTX 6000	P6000	M6000 24GB	Benefit
Architecture	Turing	Pascal	Maxwell	<i>Latest generation NVIDIA GPU technology</i>
CUDA Cores	4608	3840	3072	<i>Fast graphics and compute performance</i>
RT Cores	72	-	-	<i>GPU accelerated ray tracing for interactive and batch rendering</i>
Tensor Cores	576	-	-	<i>GPU accelerated Deep Learning for AI-augmented applications</i>
Memory	24 GB GDDR6 Up to 672 GB/s	24 GB GDDR5X Up to 432 GB/s	24 GB GDDR5 Up to 317 GB/s	<i>Smooth interaction with complex models, faster render & compute performance</i>
NVLink	2-way	-	-	<i>Scales memory & compute up to 48 GB for largest renders, models and datasets</i>
VR Ready	Multi-View Rendering	Single pass stereo	Yes	<i>Latest generation of GPU accelerated immersive VR technology</i>
VirtualLink	Yes	-	-	<i>Simplified single cable VR HMD connectivity</i>

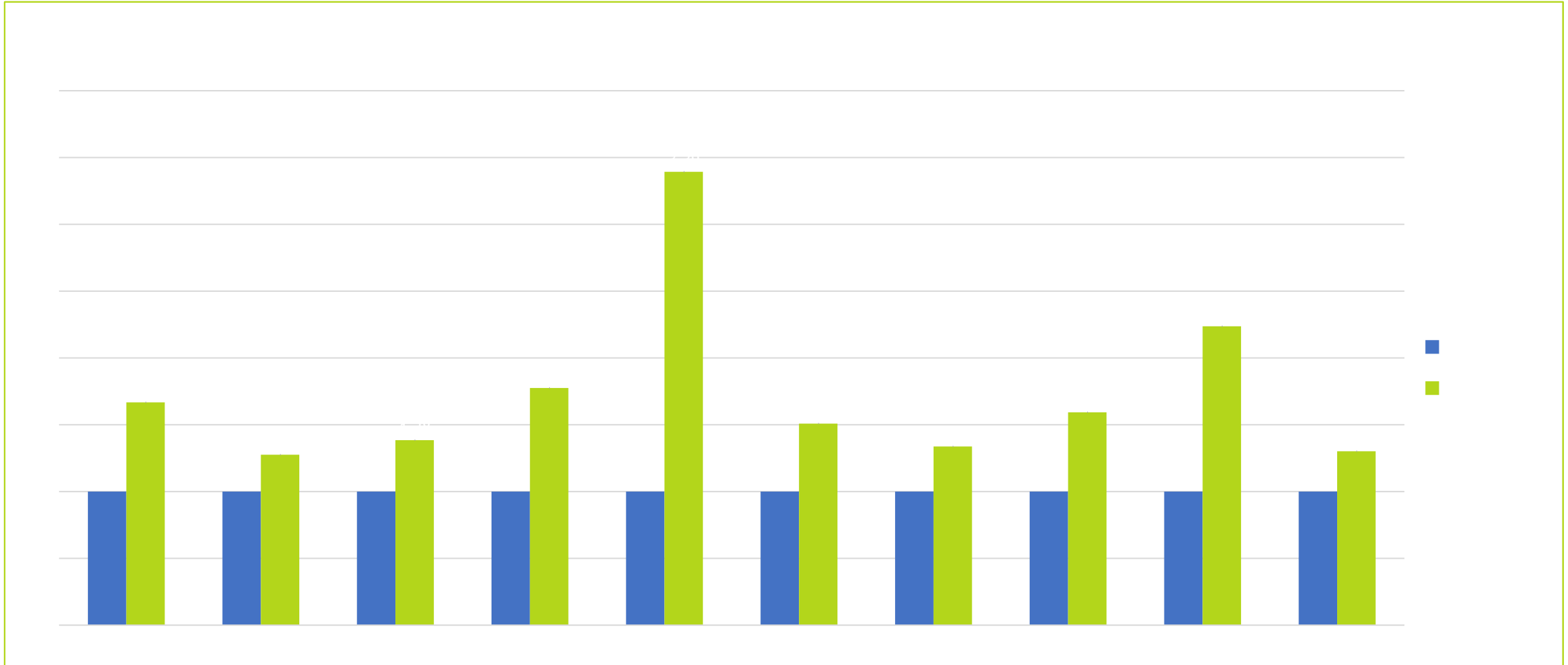
RTX 6000 UP TO 2X FASTER THAN PREVIOUS GENERATION*



Test run on a workstation with Xeon Gold 6154 3GHz (3.7 GHz turbo). 64GB RAM, Windows 10 64-bit, NVIDIA driver version 341.49 & 411.61. Performance testing completed with publicly available SPECviewperf 13 benchmark information.

*based on M6000 SPECviewperf 13 performance

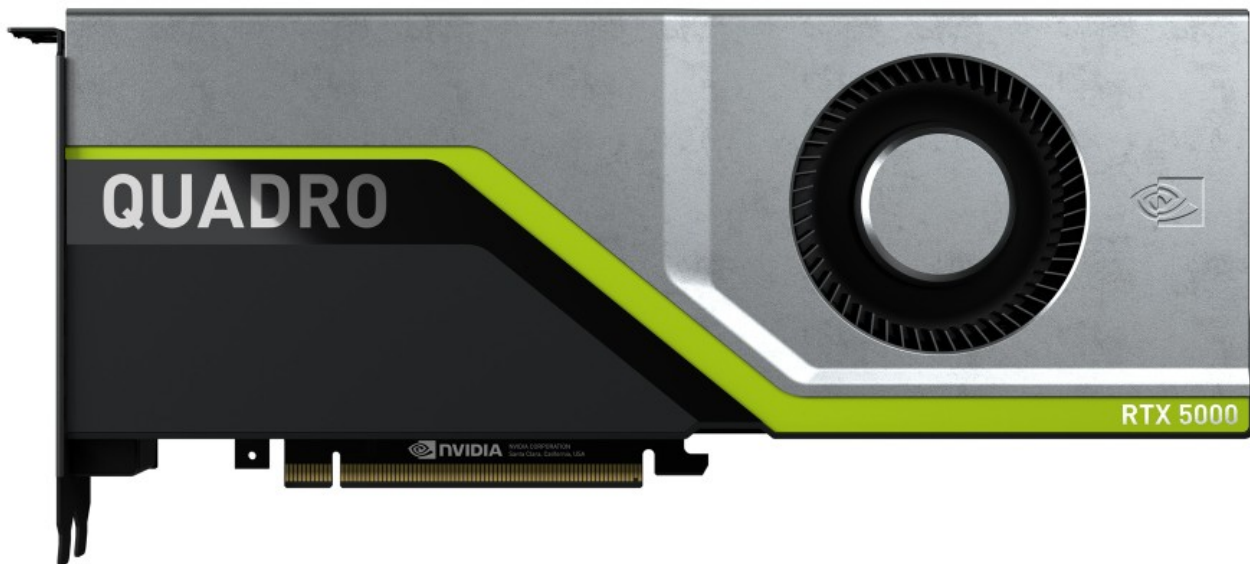
RTX 6000 MORE THAN 3X FASTER THAN COMPETITION*



Test run on a workstation with Xeon Gold 6154 3GHz (3.7 GHz turbo). 64GB RAM, Windows 10 64-bit, NVIDIA driver version 411.61, AMD driver version 18.Q4. Performance testing completed with publicly available SPECviewperf 13 benchmark information.

*based on Radeon Pro WX9100 SPECviewperf 13 performance

QUADRO RTX 5000 KEY SPECIFICATIONS

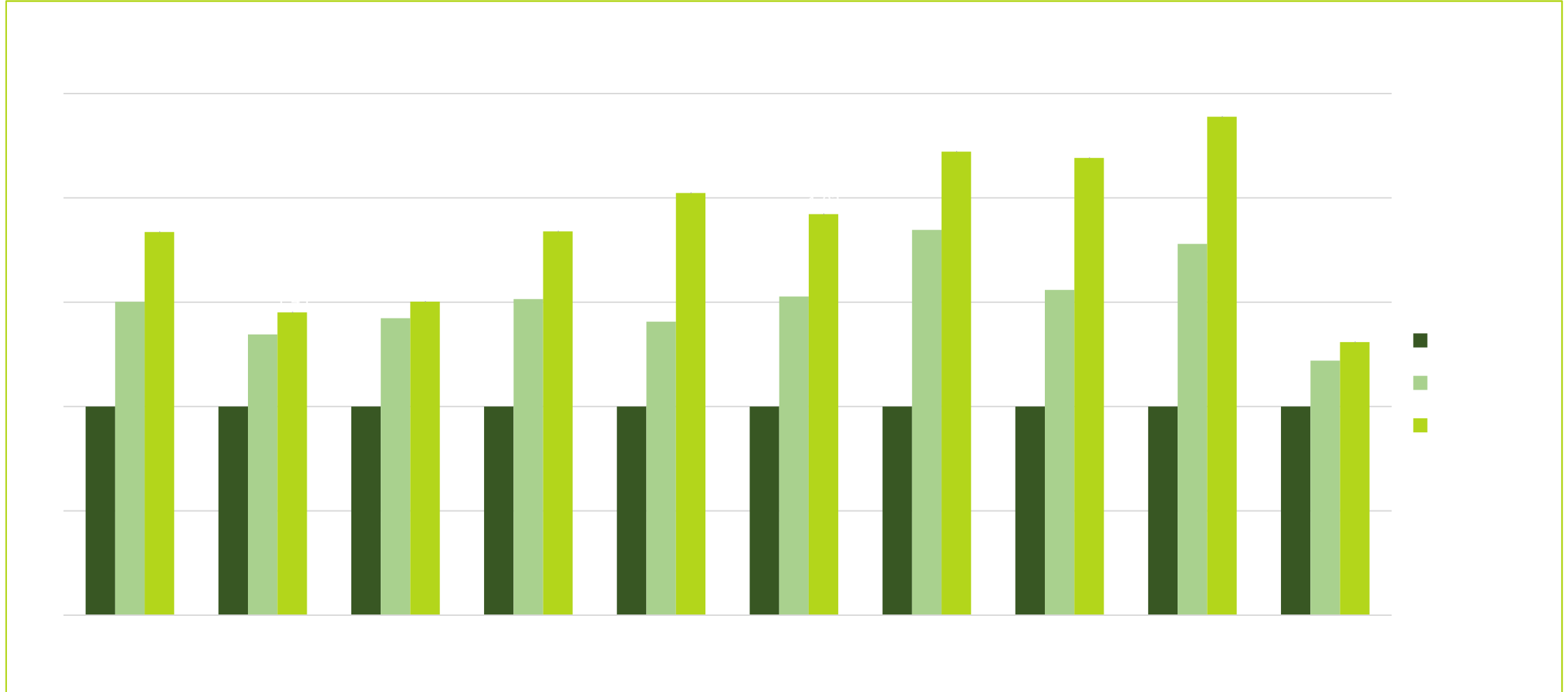


GPU Architecture	Turing
CUDA Cores	3072
RT Cores	48
Tensor Cores	384
Memory Size	16 GB GDDR6
Memory BW	Up to 448 GB/s
NVLink	2-way (2 & 3slot) 50 GB/s bidirectional
Display Support	4x DP + 1x VirtualLink
VR Ready	Yes
VirtualLink™	Yes
Advanced Display	SYNC 2
Board Power	Total Board Power: 265W Total Graphics Power: 230W
Power Connectors	1x 8-pin, 1x 6-pin PCIe

UPGRADING TO RTX 5000

	RTX 5000	P5000	M5000	Benefit
Architecture	Turing	Pascal	Maxwell	<i>Latest generation NVIDIA GPU technology</i>
CUDA Cores	3072	2560	2048	<i>Fast graphics and compute performance</i>
RT Cores	48	-	-	<i>GPU accelerated ray tracing for interactive and batch rendering</i>
Tensor Cores	384	-	-	<i>GPU accelerated Deep Learning for AI-augmented applications</i>
Memory	16GB GDDR6 Up to 448 GB/s	16 GB GDDR5X Up to 288 GB/s	8 GB GDDR5 Up to 211 GB/s	<i>Smooth interaction with complex models, faster render & compute performance</i>
NVLink	2-way	-	-	<i>Scales memory & compute up to 48 GB for largest renders, models and datasets</i>
VR Ready	Multi-View Rendering	Single pass stereo	Yes	<i>Latest generation of GPU accelerated immersive VR technology</i>
VirtualLink	Yes	-	-	<i>Simplified single cable VR HMD connectivity</i>

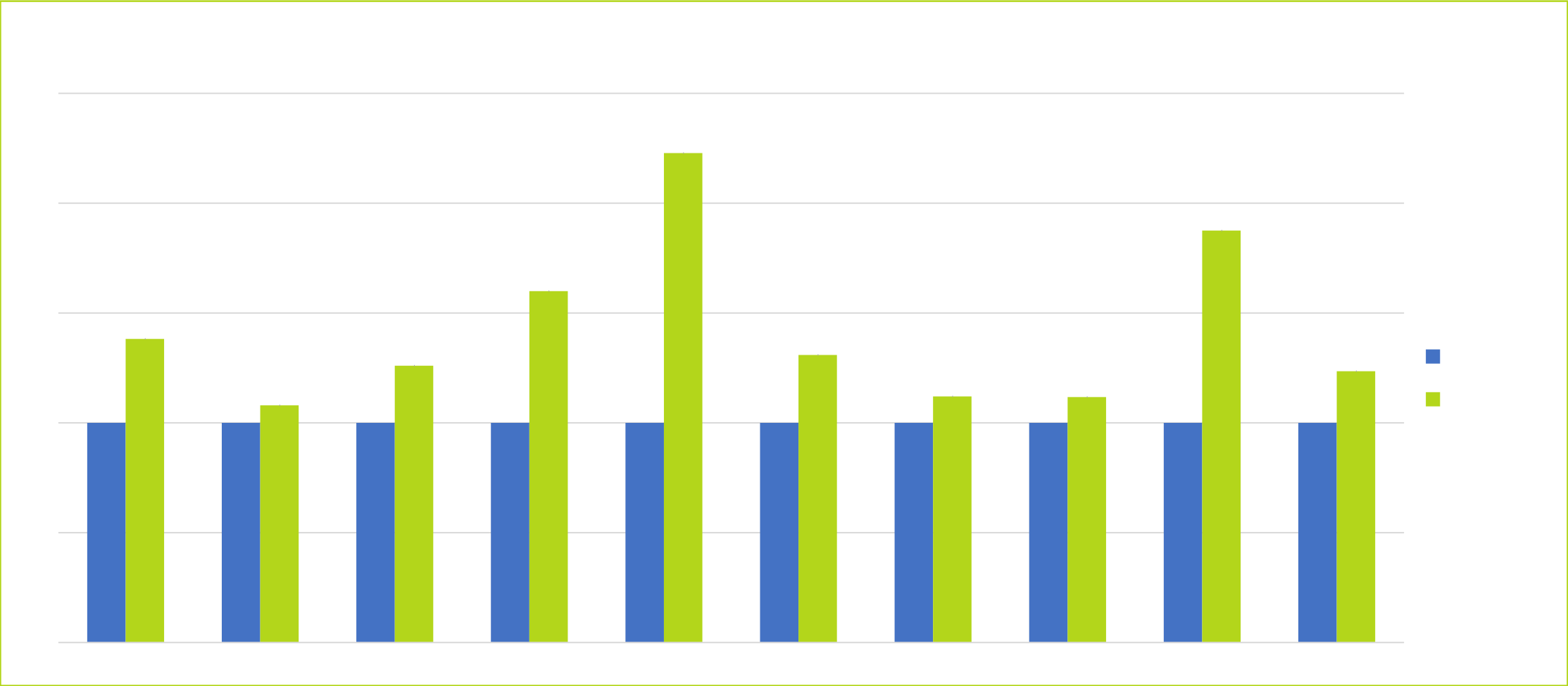
RTX 5000 MORE THAN 2X FASTER THAN PREVIOUS GENERATION*



Test run on a workstation with Xeon Gold 6154 3GHz (3.7 GHz turbo). 64GB RAM, Windows 10 64-bit, NVIDIA driver version 341.49 & 411.61. Performance testing completed with publicly available SPECviewperf 13 benchmark information.

*based on M5000 SPECviewperf 13 performance

RTX 5000 UP TO 2X FASTER THAN COMPETITION*



Test run on a workstation with Xeon Gold 6154 3GHz (3.7 GHz turbo). 64GB RAM, Windows 10 64-bit, NVIDIA driver version 411.61, AMD driver version 18.Q4. Performance testing completed with publicly available SPECviewperf 13 benchmark information.

*based on Radeon Pro WX8200 SPECviewperf 13 performance

QUADRO RTX 4000 KEY SPECIFICATIONS



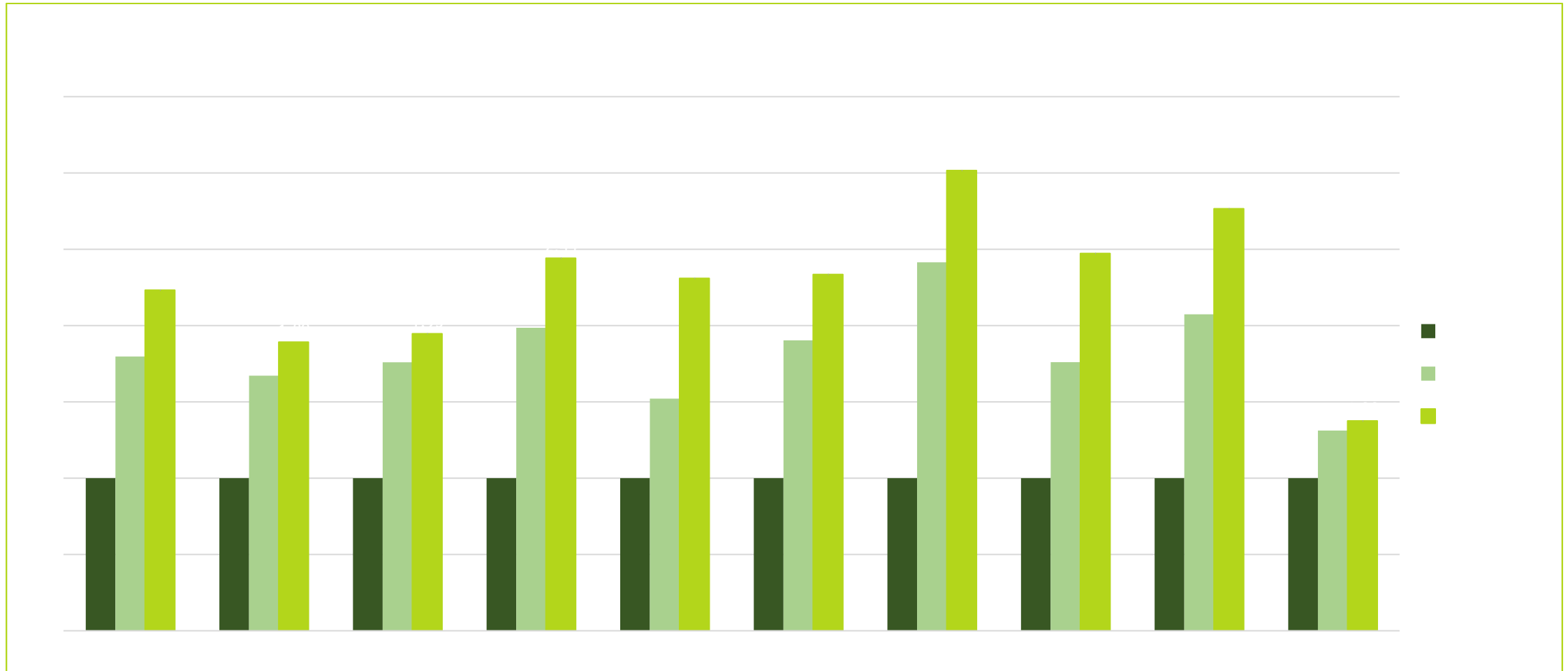
GPU Architecture	Turing
CUDA Cores	2304
RT Cores	36
Tensor Cores	288
Memory Size	8 GB GDDR6
Memory BW	Up to 416 GB/s
NVLink	N/A
Display Support*	3x DP + 1x VirtualLink
VR Ready	Yes
VirtualLink™	Yes
Advanced Display	SYNC 2
Board Power	Total Board Power: 160W Total Graphics Power: 125W
Power Connectors	1x 8-pin

*RTX 4000 can support 4x DP 1.4 using the included USB-C to DisplayPort adapter

UPGRADING TO RTX 4000

	RTX 4000	P4000	M4000	Benefit
Architecture	Turing	Pascal	Maxwell	<i>Latest generation NVIDIA GPU technology</i>
CUDA Cores	2304	1792	1664	<i>Fast graphics & compute performance</i>
RT Cores	36	-	-	<i>GPU accelerated ray tracing for interactive and batch rendering</i>
Tensor Cores	288	-	-	<i>GPU accelerated Deep Learning for AI-augmented applications</i>
Memory	8GB GDDR6 Up to 416 GB/s	8 GB GDDR5 Up to 288 GB/s	8 GB GDDR5 Up to 192 GB/s	<i>Smooth interaction with complex models, faster render & compute performance</i>
VR Ready	Multi-View Rendering	Single pass stereo	No	<i>Latest generation of GPU accelerated immersive VR technology</i>
VirtualLink	Yes	-	-	<i>Simplified single cable VR HMD connectivity</i>

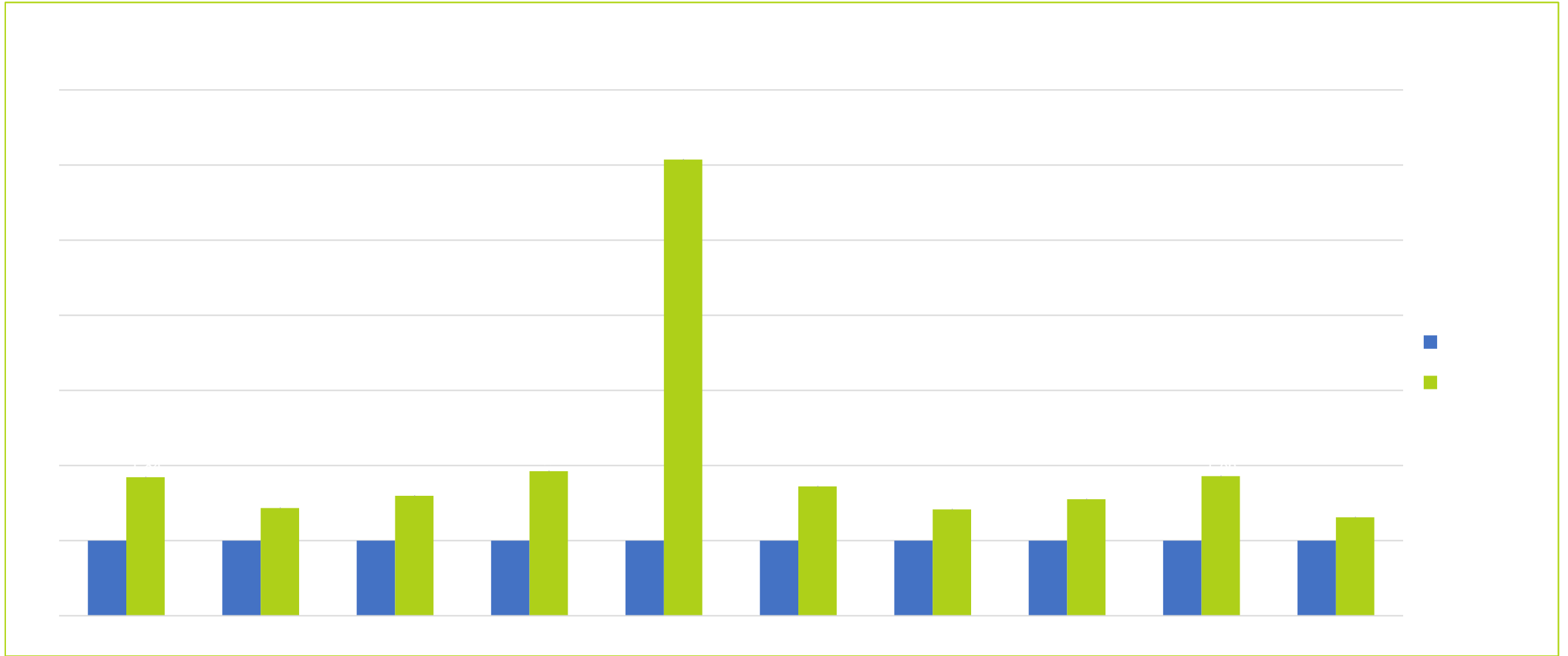
RTX 4000 UP TO 3X FASTER THAN PREVIOUS GENERATION*



Test run on a workstation with Xeon Gold 6154 3GHz (3.7 GHz turbo). 64GB RAM, Windows 10 64-bit, NVIDIA driver version 341.49 & 411.61. Performance testing completed with publicly available SPECviewperf 13 benchmark information.

*based on M4000 SPECviewperf 13 performance

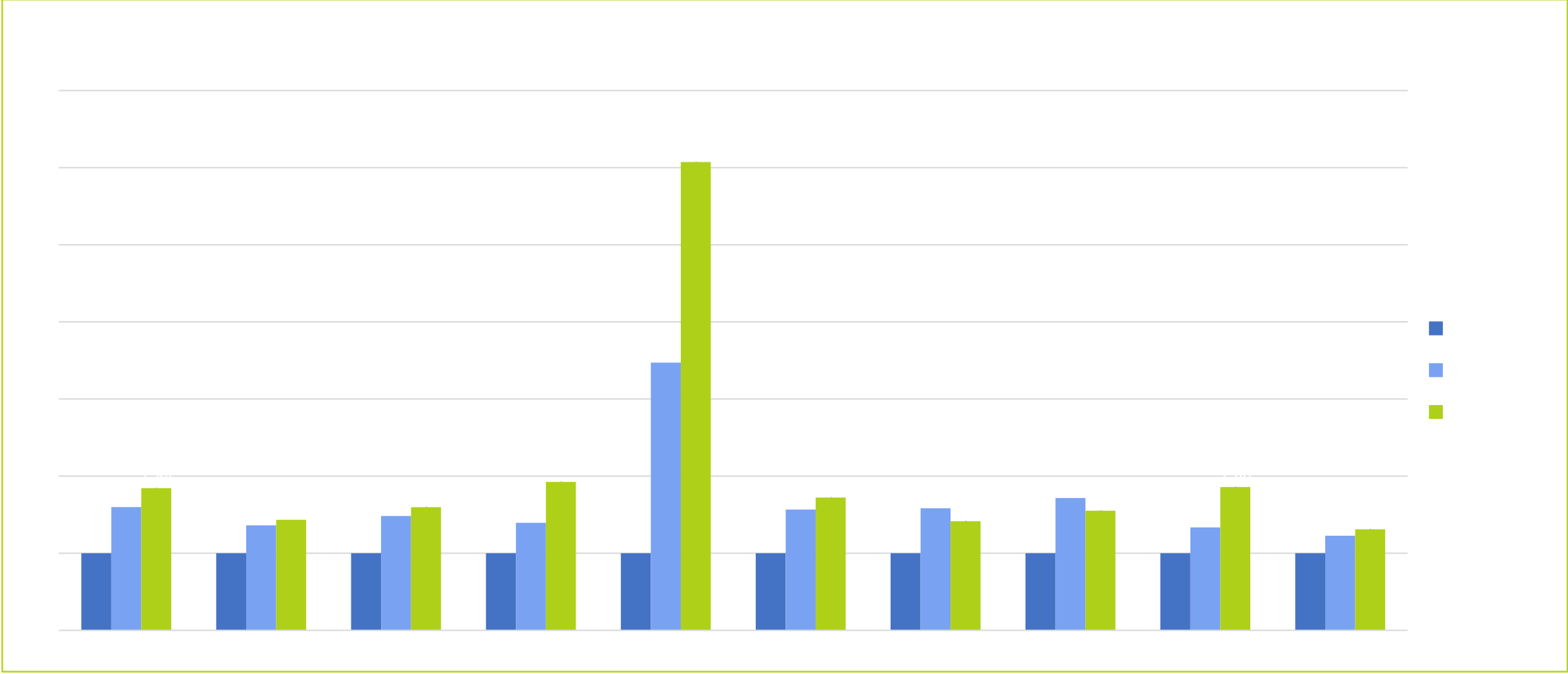
RTX 4000 1.8X FASTER THAN COMPETITION*



Test run on a workstation with Xeon Gold 6154 3GHz (3.7 GHz turbo). 64GB RAM, Windows 10 64-bit, NVIDIA driver version 411.61, AMD driver version 18.Q4. Performance testing completed with publicly available SPECviewperf 13 benchmark information.

*based on AMD WX7100 SPECviewperf 13 geomean score

RTX 4000 THE BEST PROFESSIONAL GRAPHICS CARD UNDER \$1000



Test run on a workstation with Xeon Gold 6154 3GHz (3.7 GHz turbo). 64GB RAM, Windows 10 64-bit, NVIDIA driver version 411.61, AMD driver version 18.Q4. Performance testing completed with publicly available SPECviewperf 13 benchmark information.

*based on AMD WX8200 SPECviewperf 13 geomean score

QUADRO

