ROCKY

GPU Buying Guide and FAQs

Last updated October, 2021

With Rocky 4, you can use two or more Graphic Processing Units (GPUs) to process your simulations. Before you invest in new hardware, see the FAQs below for buying guidelines and recommendations.

1. Which GPU cards are recommended for use with Rocky?

Rocky has been tested and verified with all of the following NVIDIA GPU cards:

Gaming: RTX 2080, RTX 2080 Ti, Titan RTX, RTX 3080, and RTX 3090 **PROS:** Fast when using only spherical particles, inexpensive, can be installed on individual workstations, has video output

CONS: Slow when using shaped particles

Workstation: Titan V, Quadro GP100, and Quadro GV100

PROS: Fast when using spherical and/or shaped particles, can be installed on individual workstations, has video output

CONS: More expensive

Server: Tesla P100, Tesla V100, Tesla A100, and Tesla A30

PROS: Fast when using spherical and/or shaped particles

CONS: More expensive, must be installed in a server enclosure, no video output

For best results, use only one of the above recommended GPU cards during Rocky processing.

2. There are a lot of cards on that list! How do I choose the one that is right for me?

Choosing the card that will work best for you depends upon the type of simulations you will be running, how fast you need those simulations to complete, and the budget you have to spend on your hardware.

The below table provides a quick comparison of the most common gaming and workstation cards.

	Card Name	Memory Size	Memory Bandwidth	Single Precision	Double Precision	MSRP*
Gaming Cards	GTX 1080	8GB	352 GB/s	8.2 TFlops	257 GFlops	USD 600
	GTX 1080 Ti	11GB	484 GB/s	10.6 TFlops	331 GFlops	USD 700
	RTX 2080	8GB	448 GB/s	8.9 TFlops	278 GFlops	USD 700
	RTX 2080 Ti	11GB	616 GB/s	11.7 TFlops	367 GFlops	USD 1000
	Titan RTX	24GB	672 GB/s	12.4 TFlops	388 GFlops	USD 2500
	RTX 3080	10GB	760 GB/s	25.1 TFlops	392 GFlops	USD 700
	RTX 3090	24GB	936 GB/s	29.4 TFlops	459 GFlops	USD 1500

** Last updated October, 2021*

	Card Name	Memory Size	Memory Bandwidth	Single Precision	Double Precision	MSRP*
Workstation Cards	Titan V	12GB	653 GB/s	12.2 TFlops	6144 GFlops	USD 3000
	Quadro GP100	16GB	732 GB/s	10.3 TFlops	5168 GFlops	USD 7000
	Quadro GV100	32GB	868 GB/s	14.8 TFlops	7400 GFlops	USD 9000
Server Cards	Tesla P100	16GB	730 GB/s	9.3 TFlops	4700 GFlops	USD 3000
	Tesla V100	32GB	900 GB/s	14.0 TFlops	7014 GFlops	USD 8000
	Tesla A100	40GB	1550 GB/s	19.5 TFlops	9700 GFlops	USD 9500
	Tesla A100	80GB	2000 TB/s	19.5 TFlops	9700 GFlops	–
	Tesla A30	24GB	930 TB/s	10.3 TFlops	5100 GFlops	–

** Last updated June, 2021*

3. Can you provide some examples for comparison?

Using the table above, you can see that the RTX 2080 is almost the same as the GTX 1080 (same memory and only 10% faster). However, if you take a look at the "Ti" version, then you can see a more substantial improvement with 40% more memory and performance when compared with the GTX 1080. So in this example, if you're looking for a RTX 2080, it would be beneficial to get the Ti version of it.

4. What cards are best for running only spherical particles? What about for cases using shaped particles?

Regarding particle shapes, here are some guidelines:

- When running cases with shaped particles, choosing a GPU with **double-precision performance** should be your primary focus.
- When running cases using only spherical particles, choosing a GPU with a **higher memory bandwidth** will get you better results in your processing.

It is important to note that all the 1080 and 2080 cards have fast single precision performance, but poor double precision performance. This means that they will perform very well when simulating only spherical particles, but very poorly with shaped particles. This is a critical point when you are deciding which card to acquire.

5. I have only a mid-range budget. Can you recommend a card for me?

On the mid-price range of cards, there are the Titans that come in 2 different models: RTX and V. Both have similar prices (USD ~3000) but the RTX has larger memory (24GB) and poor double precision, while the V has smaller memory (12GB) but blazing-fast double precision. Thus, you need to choose what you want: large memory (better for running cases with only spherical particles) or fast double precision (better for running cases with shaped particles).

To get both of them together (large memory and fast double precision) you would need to go for a GV100: that will be even faster than the Titan V and with more memory than the RTX, but at 3 times the cost (USD 9000).

6. If you had to recommend one, all-around best card for most situations, which would it be?

All in all, the Titan V is by far the Rocky team's preferred choice. It has a good amount of memory, blazing-fast double precision, and is not too expensive considering what it delivers in terms of processing capacity. And if it turns out your simulation does not fit onto a single-GPU, you can always use Rocky's support for multi-GPU to stack-up the GPU's combined memory.

7. Won't the (non-recommended) card I already have work just as well as a recommended one?

Different GPU cards can have an order of magnitude difference in performance, which is why we have recommended only the cards that will have the best performance on Rocky. Just because Rocky appears to run fine on a non-recommended GPU card does not mean that it is actually helping the processing performance. And if it isn't helping the performance, then there is no point to running your simulations on GPUs.

To see for yourself the huge range of performance differences, visit the links for the card types below and review the "Processing Power / Double Precision" columns of the spec tables provided: NVIDIA Tesla | NVIDIA GeForce | NVIDIA Ampere

8. Assuming I use a recommended GPU card, how much faster can I expect my simulations to run?

Compared to a CPU with 8 cores, adding even one GTX 980 has been shown to speed-up the processing time 5 fold; add in three P100s and what was once a 3-day simulation can be completed in just over an hour. But it all depends upon what you are simulating, how large your case is, and how much budget you have. See Rocky 4 with Multi-GPU: Which Hardware is Best for You? for benchmarks and speed-up comparisons.