

About Flexible GPUs

- EN
- FR

Flexible GPUs (fGPUs) are resources used for graphic rendering and large parallel computation in the Cloud. You can allocate fGPUs to your account and attach them to or detach them from your instances.

The following topics are discussed:

- [General Information](#)
- [Models of fGPUs](#)
- [Instance Type Conversion](#)

Related Pages

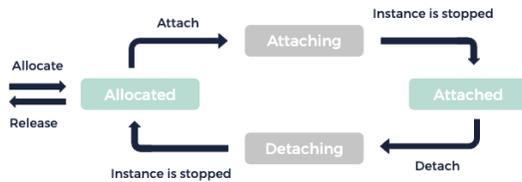
- [Flexible GPUs FAQ](#)
- [Installing an NVIDIA Driver on a Linux Instance](#)
- [Installing an NVIDIA Driver on a Windows Instance](#)

General Information

An fGPU is a resource that can be allocated to your account, and attached to or detached from your instances. You can attach several fGPUs to the same instance, provided that the fGPUs are the same models.

An fGPU can be in these different states:

- **allocated**: The fGPU is reserved for your account.
- **attaching**: The fGPU is scheduled for attachment to the specified instance. You need to stop the instance for the fGPU to reach the **attached** state.
- **attached**: The fGPU is used by the instance. It is considered a device connected to the instance.
- **detaching**: The fGPU is scheduled for detachment from its instance. You need to stop the instance for the fGPU to reach the **allocated** state.



fGPU Lifecycle



When you attach an fGPU, the instance must meet the following conditions (see the table of models of fGPUs below):

- Have less than a certain number of vCores
- Have less than a certain amount of memory
- Have a processor generation that is compatible with the fGPU

In addition, the instance must not be a dedicated instance.

Corresponding API Methods

- [FlexibleGpu](#)

Models of fGPUs

3DS OUTSCALE provides different models of fGPUs, with different amounts of video RAM (VRAM):

Model of fGPU	VRAM of the fGPU (in MiB)	Maximum Number of vCores	Maximum Amount of Memory (in GiB)	Compatible Processor Generations
nvidia-k2	4096	80	512	v3, v4
nvidia-p6	16000	80	512	v5

nvidia-p100	16000	80	512	v5
-------------	-------	----	-----	----



To make sure you have the most up-to-date values, use the [ReadFlexibleGpuCatalog](#) method as reference.

Instance Type Conversion

As with all instances, you can modify the type of an instance that has fGPUs attached to it. However, the number of attached fGPUs might be incompatible with the new instance type (number of vCores, amount of memory, and processor generation). Different behaviors occur depending on the conversion:

Old Type	New Type	Conversion Behavior
AWS	AWS	All fGPUs attached to the instance are deleted. Then as many as possible are created and attached to the converted instance.
AWS	TINA	All fGPUs attached to the instance are deleted.
TINA	AWS	All fGPUs attached to the instance are detached. Then as many as possible are created and attached to the converted instance.
TINA	TINA	If all fGPUs attached to the instance are compatible with the new instance type, the instance is converted. If not, a <code>vm-specs-mismatch</code> error is returned.

For more information on modifying instance types, see [Modifying an Instance Attribute](#).

AWS™ and **Amazon Web Services™** are trademarks of Amazon Technologies, Inc or its affiliates in the United States and/or other countries.

Windows® is a registered trademark of Microsoft Corporation in the United States and/or other countries.