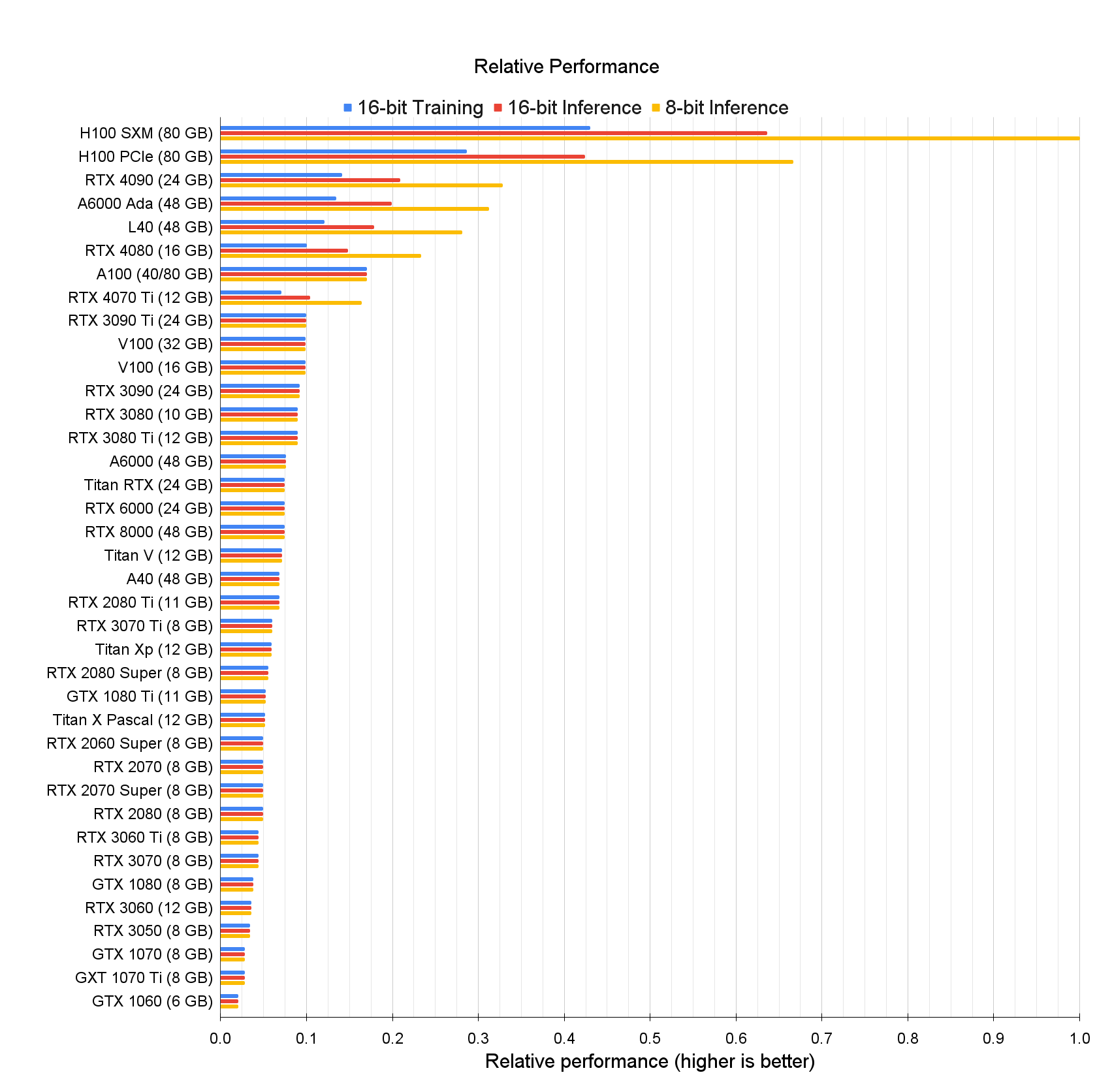
# **11/03/2023**



## AMAZON

[Amazon EC2 Instance Types - Amazon Web Services](https://aws.amazon.com/ec2/instance-types/)

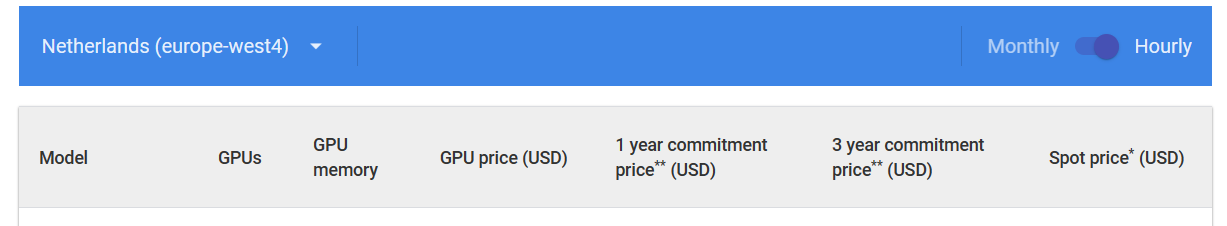
Une image contenant table

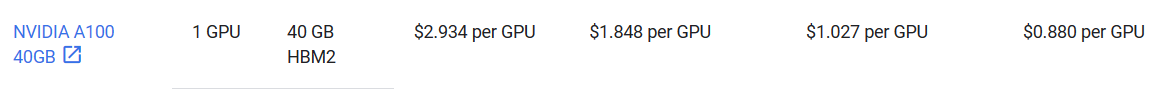
Description générée automatiquement

## GOOGLE

[GPU pricing  |  Compute Engine: Virtual Machines (VMs)  |  Google Cloud](https://cloud.google.com/compute/gpus-pricing)

[Accelerator-optimized machine family  |  Compute Engine Documentation  |  Google Cloud](https://cloud.google.com/compute/docs/accelerator-optimized-machines#a2_vms)



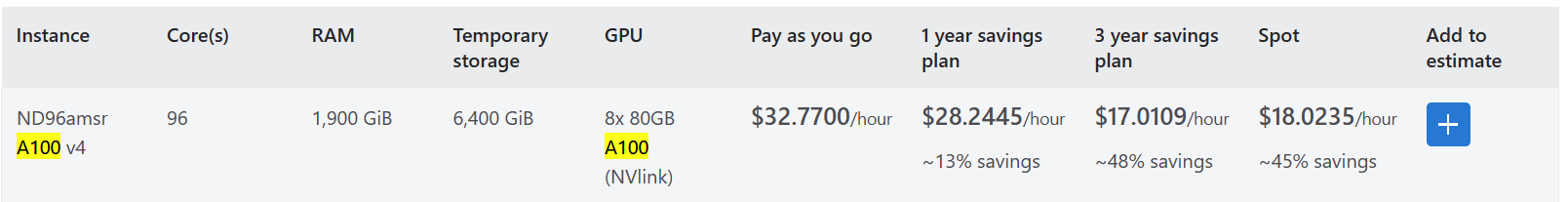


## AZURE

[Pricing - Linux Virtual Machines | Microsoft Azure](https://azure.microsoft.com/en-us/pricing/details/virtual-machines/linux/#pricing)

Une image contenant table

Description générée automatiquement



## LAMBDA

[GPU Cloud - VMs for Deep Learning | Lambda (lambdalabs.com)](https://lambdalabs.com/service/gpu-cloud#pricing)

Une image contenant table

Description générée automatiquement

## JARVIS

[Rent GPU servers for Deep Learning, AI, ML and Art generation | Jarvislabs.ai](https://jarvislabs.ai/)

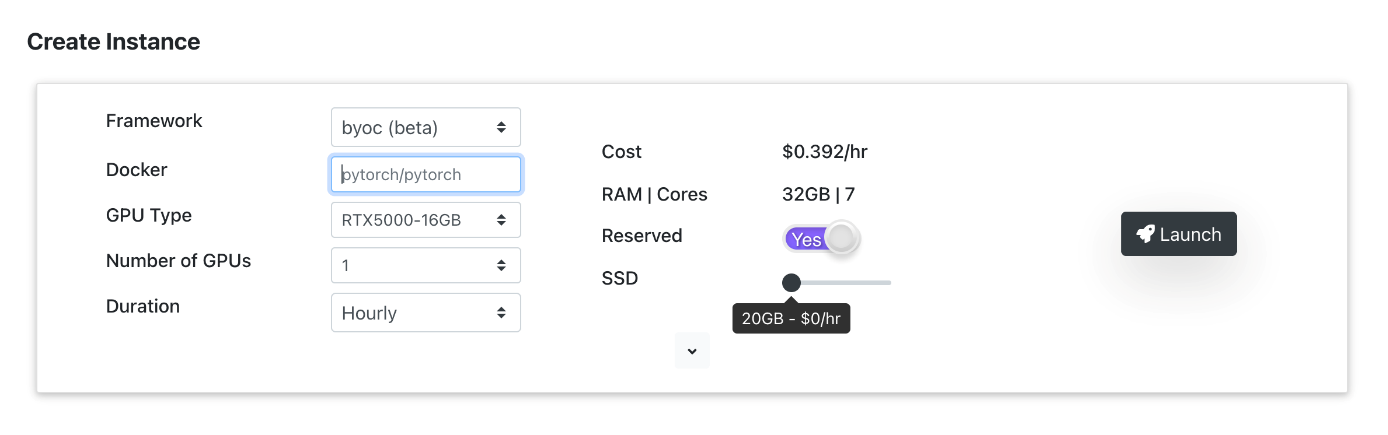
[GPU Pricing & Specifications | Jarvislabs.ai](https://jarvislabs.ai/pricing/)

Une image contenant table

Description générée automatiquement

Storage costs $0.1/GB per month and you can choose upto 500GB per instance.

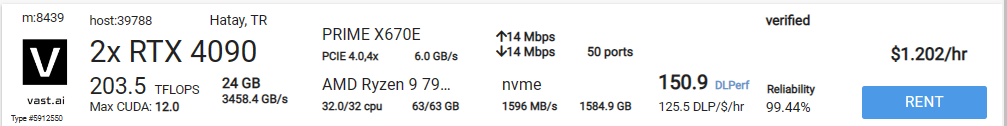
## [🚢 BYOC | Jarvislabs.ai](https://jarvislabs.ai/docs/getting-started/byoc/)

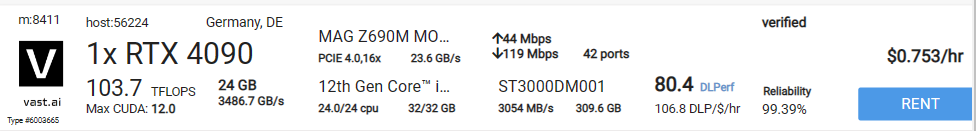


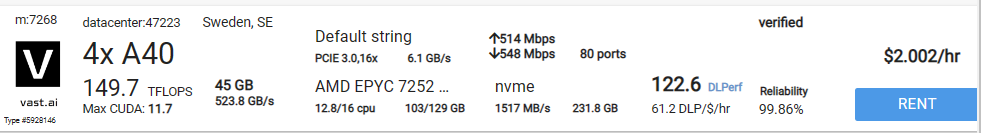
## VAST.AI

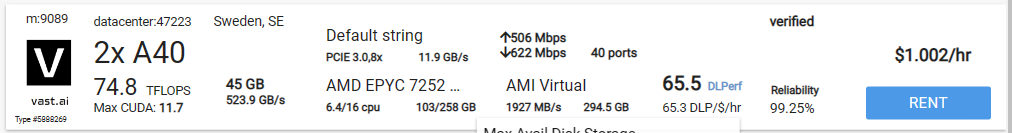
[Rent GPUs | Vast.ai](https://vast.ai/)

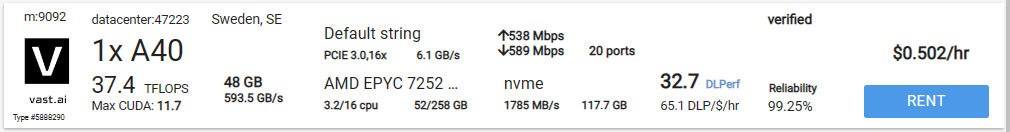
[Vast.ai | Console](https://console.vast.ai/create/)











To create an instance of type 37744 (using an ID from the search) with the vastai/tensorflow image and 32 GB of disk storage:

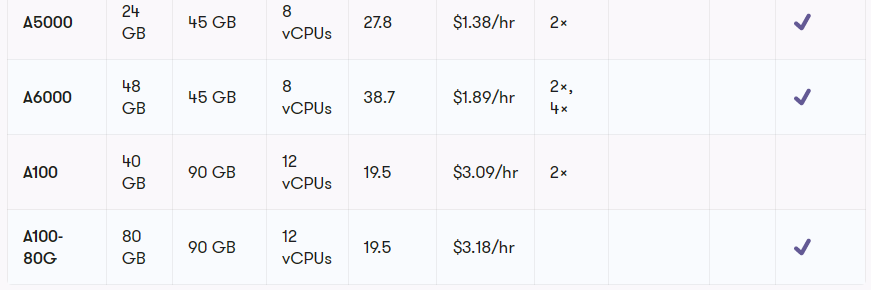
./vast create instance 36842 --image vastai/tensorflow --disk 32

## PAPERSPACE

[Paperspace - Gradient Pricing](https://www.paperspace.com/gradient/pricing)

Une image contenant texte

Description générée automatiquement



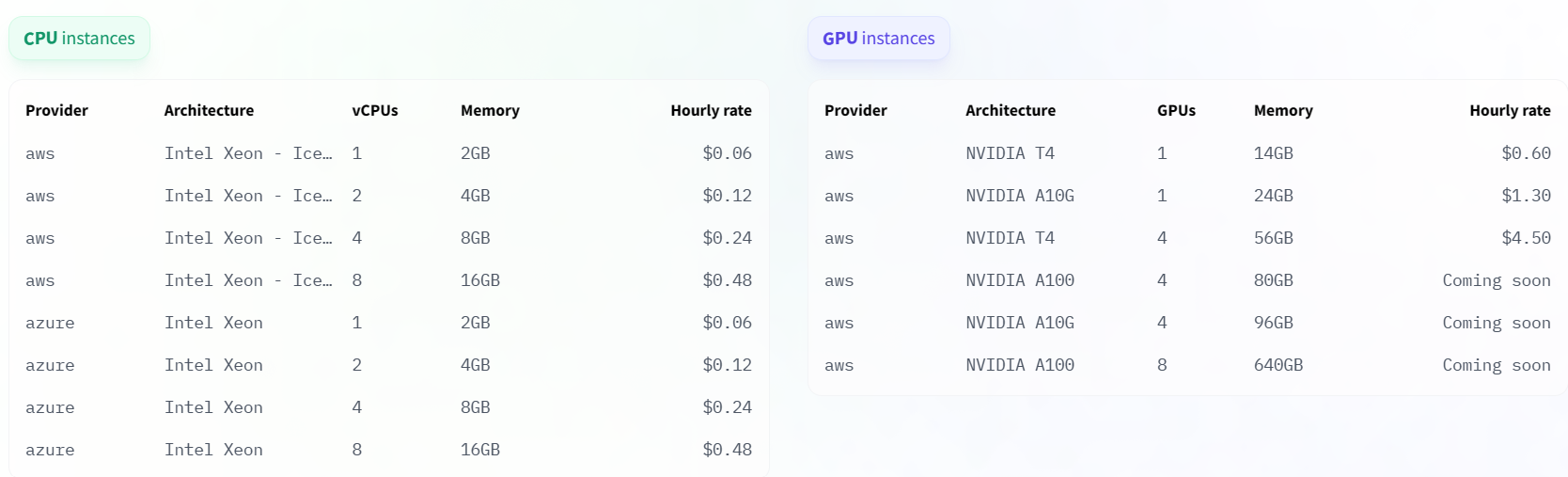
## HUGGINGFACE

Spaces

Une image contenant texte

Description générée automatiquement

Endpoints



**2019**

**AMAZON**

* **T4 / 4 vCPU / 16 Gb [120 h] : 28 $**
* **V100 / 8 vCPU / 64 Gb [40h] : 35 $**
* **200 G HDD : 10 $**
* **200 G SSD : 22 $**

<https://aws.amazon.com/ec2/instance-types/g4/>

T4 GPU

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Instance Size | vCPUs | Memory (GB) | Storage (GB) | Network Bandwidth (Gbps) | EBS Bandwidth (GBps) | On-Demand Price/hr\* |
| g4dn.xlarge | 4 | 16 | 125 | Up to 25 | Up to 3.5 | $0.526 |
| g4dn.2xlarge | 8 | 32 | 225 | Up to 25 | Up to 3.5 | $0.752 |

SPOT - Ireland

|  |  |
| --- | --- |
| g3s.xlarge | $0.2388 per Hour |
| g3.4xlarge | $0.3655 per Hour |

=> 120h training = 28 $

SPOT - US East

|  |  |
| --- | --- |
| g3s.xlarge | $0.225 per Hour |
| g3.4xlarge | $0.342 per Hour |

*À partir de l’adresse <*[*https://aws.amazon.com/ec2/spot/pricing/*](https://aws.amazon.com/ec2/spot/pricing/)*>*

V100 GPUS

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| nstance Size | GPUs - Tesla V100 | GPU Peer to Peer | GPU Memory (GB) | vCPUs | Memory (GB) | Network Bandwidth | EBS Bandwidth | On-Demand Price/hr\* |
| p3.2xlarge | 1 | N/A | 16 | 8 | 61 | Up to 10 Gbps | 1.5 Gbps | $3.06 |
| p3.8xlarge | 4 | NVLink | 64 | 32 | 244 | 10 Gbps | 7 Gbps | $12.24 |

*À partir de l’adresse <*[*https://aws.amazon.com/ec2/instance-types/p3/*](https://aws.amazon.com/ec2/instance-types/p3/)*>*

SPOT - Ireland

|  |  |
| --- | --- |
| p3.2xlarge | $1.1609 per Hour |
| p3.8xlarge | $3.966 per Hour |

=> 40h training = 35 $

SPOT - US East

|  |  |
| --- | --- |
| p3.2xlarge | $0.9519 per Hour |
| p3.8xlarge | $3.672 per Hour |

*À partir de l’adresse <*[*https://aws.amazon.com/ec2/spot/pricing/*](https://aws.amazon.com/ec2/spot/pricing/)*>*

<https://aws.amazon.com/ec2/spot/>

Run fault-tolerant workloads for up to 90% off

 While EC2 can reclaim Spot capacity with a two-minute warning, less than 5% of workloads are interrupted.

Amazon EBS General Purpose SSD (gp2) Volumes

*À partir de l’adresse <*[*https://aws.amazon.com/ebs/pricing/*](https://aws.amazon.com/ebs/pricing/)*>*

US-East : $0.1 per GB-month of provisioned storage

Ireland : $0.11

=> 200 GB = 22$/month

Amazon EBS Throughput Optimized HDD (st1) Volumes

US-East : $0.045 per GB-month of provisioned storage

Ireland : $0.05

*=> 200 GB = 10$/month*

**GOOGLE**

* **T4 / 4 vCPU / 16 Gb [120 h] : 40 $**
* **V100 / 8 vCPU / 64 Gb [40h] : 34 $**
* **200 G HDD : 9 $**
* **200 G SSD : 37 $**

Netherlands - Europe west 4

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [NVIDIA® Tesla® T4](https://www.nvidia.com/en-us/data-center/tesla-t4/) | 1 GPU | 16 GB GDDR6 | $0.95 USD per GPU | $0.29 USD per GPU | $0.600 USD per GPU | $0.430 USD per GPU |
|  | 2 GPUs | 32 GB GDDR6 |  |  |  |  |
|  | 4 GPUs | 64 GB GDDR6 |  |  |  |  |
|  |  |  |  |  |  |  |
| [NVIDIA® Tesla® V100](https://www.nvidia.com/en-us/data-center/tesla-v100/) | 1 GPU | 16 GB HBM2 | $2.55 USD per GPU | $0.74 USD per GPU | $1.606 USD per GPU | $1.147 USD per GPU |
|  | 2 GPUs | 32 GB HBM2 |  |  |  |  |
|  | 4 GPUs | 64 GB HBM2 |  |  |  |  |
|  | 8 GPUs | 128 GB HBM2 |  |  |  |  |

*À partir de l’adresse <*[*https://cloud.google.com/compute/gpus-pricing*](https://cloud.google.com/compute/gpus-pricing)*>*

GPU prices :

* 35 $
* 30 $

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Machine type** | **Virtual CPUs** | **Memory** | **Price (USD)** | **Preemptible price (USD)** |
| n1-standard-4 | 4 | 15GB | $0.2092 | $0.04410 |
| n1-standard-8 | 8 | 30GB | $0.4184 | $0.08820 |

*À partir de l’adresse <*[*https://cloud.google.com/compute/vm-instance-pricing*](https://cloud.google.com/compute/vm-instance-pricing)*>*

VM prices :

* 5 $
* 3,5 $

|  |  |
| --- | --- |
| **Type** | **Price (per GB / month)** |
| Standard provisioned space | $0.044 |
| SSD provisioned space | $0.187 |

*À partir de l’adresse <*[*https://cloud.google.com/compute/disks-image-pricing*](https://cloud.google.com/compute/disks-image-pricing)*>*

Disk prices

* 9 $
* 37,4 $

**MICROSOFT**

* **T4 / 4 vCPU / 16 Gb [120 h] :**
* **V100 / 8 vCPU / 64 Gb [40h] : 153 $ (not preemptible)**
* **200 G HDD : 11.3 $**
* **200 G SSD : 19.2 $**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Size** | **vCPU** | **Memory: GiB** | **Temp storage (SSD): GiB** | **GPU** | **GPU memory: GiB** | **Max data disks** | **Max uncached disk throughput: IOPS / MBps** | **Max NICs** |
| Standard\_NC6s\_v3 | 6 | 112 | 736 | 1 | 16 | 12 | 20000 / 200 | 4 |
| Standard\_NC12s\_v3 | 12 | 224 | 1474 | 2 | 32 | 24 | 40000 / 400 | 8 |

*À partir de l’adresse <*[*https://docs.microsoft.com/en-us/azure/virtual-machines/windows/sizes-gpu*](https://docs.microsoft.com/en-us/azure/virtual-machines/windows/sizes-gpu)*>*

NC6 west europe = $3.823/hour

*À partir de l’adresse <*[*https://azure.microsoft.com/en-us/pricing/details/virtual-machines/linux/*](https://azure.microsoft.com/en-us/pricing/details/virtual-machines/linux/)*>*

**Paperspace**

**V100 [40h] / 8 vCPU / 200 G : 8 + 40 \* 1.15 = 54 $**

G1 Developer : 8$ / month

Free + Low-Mid Instance Types

200GB Persistent Storage

Auto-shutdown (Configurable)

Email Support

Unlimited Jobs (2 concurrent)

10 Notebooks Limit (5 concurrent)

V100

$1.15/ hour

16GB GDDR5 DEDICATED

30GB RAM

8 vCPU

900 GB/s memory bandwidth

5,120 CUDA cores