

GPU 云服务器

实例类型



腾讯云

【 版权声明 】

©2013–2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

文档目录

实例类型

计算型实例

渲染型实例

实例类型

计算型实例

最近更新时间：2025-06-09 18:06:32

GPU 计算型实例能够提供强大的计算能力，从容应对高实时、高并发的海量计算场景，广泛适用于深度学习、科学计算等 GPU 通用计算场景。腾讯云 GPU 云服务器以和 **云服务器 CVM** 一致的管理方式，提供快速、稳定、弹性的计算服务。

说明：
若您的 GPU 实例用于 3D 图形渲染任务，则建议您使用已配置 vDWS/vWS License 并安装 GRID driver 的 **渲染型实例**，以此免除手动配置 GPU 图形图像处理基础环境步骤。

计算型实例总览

GPU 云服务器计算型系列提供以下实例：

售卖情况	实例	GPU 类型	售卖地域
主售	PNV 4	NVIDIA A10	广州、上海、南京、北京
	GT4	NVIDIA A100	广州、上海、南京、北京
	GN10 Xp	NVIDIA V100	广州、上海、南京、北京、成都、重庆、新加坡、曼谷、首尔、东京、法兰克福、硅谷
	GN7	NVIDIA T4	广州、上海、南京、北京、成都、重庆、中国香港、新加坡、曼谷、雅加达、首尔、东京、法兰克福、硅谷、弗吉尼亚、圣保罗
	GN7v i	NVIDIA T4	上海、南京
	PTX1	腾讯紫霄 C100	广州、上海、南京
邀测	PNV 5b	NVIDIA GPU	北京、上海、南京、广州、新加坡、法兰克福
	PNV 6	NVIDIA GPU	-

在售	GI3X	NVIDIA T4	广州、上海、南京、北京、成都、重庆
	GN10X	NVIDIA V100	广州、上海、南京、北京、成都、重庆、新加坡、法兰克福、硅谷
	GN8	NVIDIA P40	广州、上海、北京、成都、重庆、中国香港、硅谷
	GN6 GN6S	NVIDIA P4	- GN6：成都 - GN6S：广州、上海、北京

计算型实例选型推荐

腾讯云提供了类型丰富的 GPU 计算实例，可满足不同业务应用场景的需求。请参考下表，并结合实际需求选择合适的计算实例。

GPU 云服务器计算型实例选型推荐如下表，其中 ✓ 为支持，★ 为推荐。

功能\实例	PN V4	G T4	GN10 Xp	G N 7	GN 7vi	GI 3 X	GN1 0X	G N 8	GN6/ GN6S	P T X1
图形图像处理	✓	—	✓	✓	✓	✓	✓	✓	✓	—
视频编解码	✓	—	✓	★	★	★	✓	✓	✓	—
深度学习训练	✓	★	★	✓	✓	✓	★	★	✓	—
深度学习推理	★	✓	★	★	★	★	★	✓	✓	★
科学计算	—	★	★	—	—	—	★	—	—	—

 **注意：**

- 以上推荐用途仅供参考，请根据实际需要进行选择。
- NVIDIA 系列 GPU 实例如用作通用计算，则需安装 Tesla Driver + CUDA，安装方法请参见 [安装 NVIDIA Tesla 驱动指引](#) 和 [安装 CUDA 驱动指引](#)。
- NVIDIA 系列 GPU 实例如用作 3D 图形渲染任务（高性能图形处理，视频编解码等），则需安装 GRID Driver 和配置 License Server，安装方法请参见 [安装 NVIDIA GRID 驱动](#)。

实例规格

计算型 PNV4

计算型 PNV4 不仅适用于深度学习等 GPU 通用计算场景，也适用于图形图像处理（3D 渲染，视频编解码）场景。

适用场景

性价比高，适用于如下场景：

- 深度学习的推理场景和小规模训练场景。例如：
 - 大规模部署的 AI 推理
 - 深度学习小规模训练
- 图形图像处理场景。例如：
 - 图形图像处理
 - 视频编解码
 - 图形数据库

硬件规格

- **GPU:** NVIDIA® A10 (FP32 31.2 TFLOPS, TF32 62.5 TFLOPS, FP16 125 TFLOPS, INT8 250 TOPS)。
- **CPU:** 2.55GHz AMD EPYC™ Milan 处理器，睿频3.5GHz。
- **内存:** 搭配八通道 DDR4 内存。
- **存储:** 可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络:** 默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
PNV4.7XLARGE116	NVIDIA A10 * 1	24GB * 1	28	116	13	230万	28
PNV4.14XLARGE232	NVIDIA A10 * 2	24GB * 2	56	232	25	470万	48
PNV4.28XLARGE466	NVIDIA A10 * 4	24GB * 4	112	466	50	950万	48
PNV4.56XLA	NVIDIA	24GB	224	932	100	1900万	48

RGE932	A10 * 8	* 8					
--------	---------	-----	--	--	--	--	--

计算型 GT4

计算型 GT4 适用于深度学习、科学计算等 GPU 通用计算场景。

适用场景

GT4 具有强大的双精度浮点运算能力，适用于大规模深度学习训练、推理和科学计算场景。例如：

- 深度学习
- 高性能数据库
- 计算流体力学
- 计算金融
- 地震分析
- 分子建模
- 基因组学及其他

硬件规格

- GPU:** NVIDIA® A100 NVLink 40GB (FP64 9.7 TFLOPS, FP32 19.5 TFLOPS, 600GB/s NVLink)。
- CPU:** 2.6GHz AMD EPYC™ ROME 处理器，睿频3.3GHz。
- 内存:** 搭配八通道 DDR4 内存。
- 存储:** 可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- 网络:** 最高可支持50Gbps内网带宽，超高网络收发包能力，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽 (Gbps)	网络收发包(PPS)	队列数
GT4.4XLAR GE96	NVIDIA A100 * 1	40GB * 1	16	96	5	120万	4
GT4.8XLAR GE192	NVIDIA A100 * 2	40GB * 2	32	192	10	235万	8
GT4.20XLAR GE474	NVIDIA A100 * 4	40GB * 4	82	474	25	600万	16
GT4.41XLAR	NVIDIA	40GB	164	948	50	1200万	32

GE948

A100 * 8

* 8

说明：

GPU 驱动：NVIDIA A100 系列需要安装 NVIDIA Tesla 450 版本以上驱动，驱动版本信息请参见 [NVIDIA 官方文档](#)。

计算型 GN10Xp

计算型 GN10Xp 不仅适用于深度学习、科学计算等 GPU 通用计算场景，也适用于图形图像处理（3D 渲染，视频编解码）场景。

适用场景

GN10Xp 具有强大的双精度浮点运算能力，适用于如下场景：

- 大规模深度学习训练，推理和科学计算场景。例如：

- 深度学习
- 高性能数据库
- 计算流体动力学
- 计算金融
- 地震分析
- 分子建模
- 基因组学及其他

- 图形图像处理场景。例如：

- 图形图像处理
- 视频编解码
- 图形数据库

硬件规格

- CPU：**GN10Xp 配置 Intel® Xeon® Platinum 8255C CPU，主频2.5GHz。
- GPU：**NVIDIA® Tesla® V100 NVLink 32GB (15.7TFLOPS 单精度浮点计算，7.8TFLOPS 双精度浮点计算，125TFLOPS Tensor Core 深度学习加速，300GB/s NVLink)。
- 内存：**DDR4，内存速率达2666MT/s。
- 存储：**可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- 网络：**默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vC PU	内存 (Gi B)	内网带 宽	网络收 发包 (PPS)	队列数
----	-----	-----------	----------	-----------------	----------	--------------------	-----

					(Gbps)		
GN10Xp.2XLA RGE40	NVIDIA V100 * 1	32GB * 1	10	40	3	80万	2
GN10Xp.5XLA RGE80	NVIDIA V100 * 2	32GB * 2	20	80	6	150万	5
GN10Xp.10XLA RGE160	NVIDIA V100 * 4	32GB * 4	40	160	12	250万	10
GN10Xp.20XLARGE320	NVIDIA V100 * 8	32GB * 8	80	320	24	490万	16

计算型 GN7

NVIDIA 实例 GN7 不仅适用于深度学习等 GPU 通用计算场景，也适用于图形图像处理（3D 渲染，视频编解码）场景。

适用场景

性价比高，适用于如下场景：

- 深度学习的推理场景和小规模训练场景。例如：
 - 大规模部署的 AI 推理
 - 深度学习小规模训练
- 图形图像处理场景。例如：
 - 图形图像处理
 - 视频编解码
 - 图形数据库

硬件规格

- **CPU**：Intel® Xeon® Platinum 8255C CPU，主频 2.5 GHz。
- **GPU**：NVIDIA® Tesla® T4（8.1 TFLOPS 单精度浮点计算，130 INT8 TOPS，260 INT4 TOPS）。
- **内存**：DDR4，内存速率达2666MT/s。
- **存储**：可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络**：默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
----	-----	--------	------	----------	-------------	-------------	-----

GN7.2XLARGE 32	NVIDIA T4 * 1	16GB * 1	8	32	3	60万	8
GN7.5XLARGE 80	NVIDIA T4 * 1	16GB * 1	20	80	7	140万	10
GN7.8XLARGE 128	NVIDIA T4 * 1	16GB * 1	32	128	10	240万	16
GN7.10XLARG E160	NVIDIA T4 * 2	16GB * 2	40	160	13	280万	20
GN7.20XLARG E320	NVIDIA T4 * 4	16GB * 4	80	320	25	560万	32

视频增强型 GN7vi

NVIDIA 实例 GN7vi 是在 GN7 基础上配置腾讯自研的明眸视频融合 AI 技术，包括极速高清编解码引擎和画质增强工具包，适用于点播、直播场景。使用该实例，您可在实例内部使用腾讯云自研的极速高清编解码和 AI画质增强功能。

 **说明：**

如果您希望了解更多 GPU 视频增强型 GN7vi 产品信息，请前往 [使用咨询](#) 页面提供联系方式，我们将会
有专人联系您。

硬件规格

- **CPU：**Intel® Xeon® Platinum 8255C CPU，主频2.5GHz。
- **GPU：**NVIDIA® Tesla® T4（8.1 TFLOPS 单精度浮点计算，130 INT8 TOPS，260 INT4 TOPS）。
- **内存：**DDR4，内存速率达2666MT/s。
- **存储：**可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络：**默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vC PU	内 存 (Gi B)	内网带 宽 (Gbp s)	网络收 发包 (PPS)	队列数
GN7vi.5XLARG E80	NVIDIA T4 * 1	16GB * 1	20	80	6	140万	20
GN7vi.10XLAR GE160	NVIDIA T4 * 2	16GB * 2	40	16 0	13	280万	32

GN7vi.20XLAR GE320	NVIDIA T4 * 4	16GB * 4	80	32 0	25	560万	32
-----------------------	------------------	-------------	----	---------	----	------	----

计算型 PNV5b（邀测）

计算型 PNV5b 采用全新架构的 GPU 计算卡，48GB GDDR6显存容量，支持 FP32、FP16、BF16、FP8、INT8等计算格式，搭配 AMD EPYC™ Genoa 处理器，适用于深度学习推理、广告推荐等场景，也适用于图形图像处理（3D 渲染，视频编解码）场景。

说明：
该实例暂时处于白名单邀测状态，请联系您的售前经理开通实例购买权限。

适用场景

- 深度学习的推理场景和广告推荐场景。例如：
 - 大语言模型推理
 - 广告推荐
 - AIGC 图像、视频生成
- 图形图像处理场景。例如：
 - 图形图像处理
 - 视频编解码
 - 图形数据库

硬件规格

- **CPU：**AMD EPYC™ Genoa 处理器，睿频3.7GHz。
- **内存：**搭配十二通道 DDR5 内存。
- **存储：**可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络：**默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vCP U	内存 (Gi B)	内网带 宽 (Gbp s)	网络收发 包 (PPS)	队列 数
PNV5b.8XLARG E96	NVIDIA GPU * 1	48GB * 1	32	96	13	560万	32
PNV5b.12XLAR GE192	NVIDIA GPU * 1	48GB * 1	48	192	13	560万	48

PNV5b.16XLAR GE192	NVIDIA GPU * 2	48GB * 2	64	192	25	1120万	48
PNV5b.24XLAR GE384	NVIDIA GPU * 2	48GB * 2	96	384	25	1120万	48
PNV5b.32XLAR GE384	NVIDIA GPU * 4	48GB * 4	128	384	50	2250万	48
PNV5b.48XLAR GE768	NVIDIA GPU * 4	48GB * 4	192	768	50	2250万	48
PNV5b.64XLAR GE768	NVIDIA GPU * 8	48GB * 8	256	768	100	4500万	48
PNV5b.96XLAR GE1536	NVIDIA GPU * 8	48GB * 8	384	1536	100	4500万	48

计算型 PNV6（邀测）

计算型 PNV6 适用于深度学习推理及小规模训练场景。

 **说明：**
该实例暂时处于白名单邀测状态，请联系您的售前经理开通实例购买权限。

适用场景

性价比高，适用于如下场景：

- 深度学习的推理场景和小规模训练场景。例如：
 - 大语言模型推理
 - 广告推荐
 - 自动驾驶
 - 计算机视觉

硬件规格

- **CPU：**AMD EPYC™ Genoa 处理器，睿频3.7GHz。
- **内存：**搭配十二通道 DDR5 内存。
- **存储：**可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络：**默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	vCP U	内存	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
----	-----	----------	----	----------------	----------------	-----

			(GiB)			
PNV6.4XLAR GE160	NVIDIA GPU * 1	16	160	8	180万	16
PNV6.8XLAR GE320	NVIDIA GPU * 2	32	320	13	370万	32
PNV6.16XLAR GE640	NVIDIA GPU * 4	64	640	25	750万	48
PNV6.32XLA RGE1280	NVIDIA GPU * 8	128	1280	50	1500万	48
PNV6.96XLA RGE2304	NVIDIA GPU * 8	384	2304	100	4500万	48

推理型 GI3X

NVIDIA 实例 GI3X 适用于深度学习等 GPU 通用计算场景，也适用于图形图像处理（3D 渲染，视频编解码）场景。

适用场景

性价比高，适用于如下场景：

- 深度学习的推理场景和小规模训练场景。例如：
 - 大规模部署的 AI 推理
 - 深度学习小规模训练
- 图形图像处理场景。例如：
 - 图形图像处理
 - 视频编解码
 - 图形数据库

硬件规格

- **CPU**：2.6GHz AMD EPYC™ ROME 处理器，睿频3.3GHz。
- **GPU**：NVIDIA® Tesla® T4（8.1 TFLOPS 单精度浮点计算，130 INT8 TOPS，260 INT4 TOPS）。
- **内存**：搭配八通道 DDR4，内存计算性能稳定。
- **存储**：可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络**：默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

GI3X 实例提供以下配置：

规格	GPU	GPU 显存	vCP U	内 存 (Gi B)	内网带 宽 (Gbp s)	网络收 发包 (PPS)	队列数
GI3X.8XLARG E64	NVIDIA T4 * 1	16GB * 1	32	64	5	140万	8
GI3X.22XLAR GE226	NVIDIA T4 * 2	16GB * 2	90	22 6	13	375万	16
GI3X.45XLAR GE452	NVIDIA T4 * 4	16GB * 4	180	45 2	25	750万	32

计算型 GN10X

计算型 GN10X 不仅适用于深度学习、科学计算等 GPU 通用计算场景，也适用于图形图像处理（3D 渲染，视频编解码）场景。

适用场景

GN10X具有强大的双精度浮点运算能力，适用于如下场景：

- 大规模深度学习训练，推理和科学计算场景。例如：

- 深度学习
- 高性能数据库
- 计算流体动力学
- 计算金融
- 地震分析
- 分子建模
- 基因组学及其他

- 图形图像处理场景。例如：

- 图形图像处理
- 视频编解码
- 图形数据库

硬件规格

- CPU:** GN10X 配置 Intel® Xeon® Gold 6133 CPU，主频2.5GHz。
- GPU:** NVIDIA® Tesla® V100 NVLink 32GB（15.7TFLOPS 单精度浮点计算，7.8TFLOPS 双精度浮点计算，125TFLOPS Tensor Core 深度学习加速，300GB/s NVLink）。
- 内存:** DDR4，内存速率达2666MT/s。

- **存储**：可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络**：默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

GN10X 实例提供以下配置：

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
GN10X.2XLA RGE40	NVIDIA V100 * 1	32GB * 1	8	40	3	80万	2
GN10X.4XLA RGE80	NVIDIA V100 * 2	32GB * 2	18	80	7	150万	4
GN10X.9XLA RGE160	NVIDIA V100 * 4	32GB * 4	36	160	13	250万	9
GN10X.18XLA RGE320	NVIDIA V100 * 8	32GB * 8	72	320	25	490万	16

计算型 GN8

NVIDIA 实例 GN8 不仅适用于深度学习等 GPU 通用计算场景，也适用于图形图像处理（3D 渲染，视频编解码）场景。

适用场景

适用于如下场景：

- 深度学习的推理和训练场景。例如：
 - 大吞吐量的 AI 推理
 - 深度学习
- 图形图像处理场景。例如：
 - 图形图像处理
 - 视频编解码
 - 图形数据库

硬件规格

- **CPU**：Intel® Xeon® E5-2680 v4 CPU，主频2.4GHz。
- **GPU**：NVIDIA® Tesla® P40（12TFLOPS 单精度浮点计算，47INT8 TOPS）。
- **内存**：DDR4，内存速率达2666MT/s。
- **存储**：可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。

- **网络**：默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

GN8 实例提供以下配置：

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
GN8.LARGE56	NVIDIA P40 * 1	24GB * 1	6	56	1.5	45万	8
GN8.3XLARGE112	NVIDIA P40 * 2	24GB * 2	14	112	2.5	50万	8
GN8.7XLARGE224	NVIDIA P40 * 4	24GB * 4	28	224	5	70万	14
GN8.14XLARGE448	NVIDIA P40 * 8	24GB * 8	56	448	10	70万	28

计算型 GN6/GN6S

NVIDIA 实例 GN6/GN6S 不仅适用于深度学习等 GPU 通用计算场景，也适用于图形图像处理（3D 渲染，视频编解码）场景。

适用场景

性价比高，适用于如下场景：

- 深度学习的推理场景和小规模训练场景。例如：
 - 大规模部署的 AI 推理
 - 深度学习小规模训练
- 图形图像处理场景。例如：
 - 图形图像处理
 - 视频编解码
 - 图形数据库

硬件规格

- **CPU**：GN6 配置 Intel® Xeon® E5-2680 v4 CPU，主频2.4GHz。GN6S 配置 Intel® Xeon® Silver 4110 CPU，主频2.1GHz。
- **GPU**：NVIDIA® Tesla® P4（5.5TFLOPS 单精度浮点计算，22INT8 TOPS）。
- **内存**：DDR4，内存速率达2666MT/s。
- **存储**：可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。

- **网络**：默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

GN6/GN6S 实例提供以下配置：

规格	GPU	GPU 显存	vCPU U	内存 (Gi B)	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
GN6.7XLARGE48	NVIDIA P4 * 1	8GB * 1	28	48	5	120万	14
GN6.14XLARGE96	NVIDIA P4 * 2	8GB * 2	56	96	10	120万	28
GN6S.LARGE20	NVIDIA P4 * 1	8GB * 1	4	20	5	50万	8
GN6S.2XLARGE40	NVIDIA P4 * 2	8GB * 2	8	40	9	80万	8

NPU 计算型 PTX1

PTX1 适用于深度学习推理计算，在 CV、OCR、ASR 等场景有良好的性能表现。

 **说明：**
该实例暂时处于白名单开放状态，请您通过 [售前在线咨询](#) 进行实例购买权限开通。

适用场景

- 性价比高，适用于深度学习推理场景。例如：
 - 大规模部署的 AI 推理
 - 图像识别
 - 文字识别
 - 语音识别

硬件规格

- **CPU**：2.55GHz AMD EPYCTM Milan 处理器，睿频3.5GHz。
- **NPU**：腾讯紫霄C100（120 TFLOPS FP16）
- **存储**：可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络**：最高可支持100Gbps内网带宽，超高网络收发包能力，实例网络性能与规格对应。[公网网络](#) 可按需配置。

PTX1 实例提供以下配置：

规格	GPU	GPU 显存	vC PU	内 存 (Gi B)	内网 带宽 (Gbp s)	网络收 发包 (PPS)	队列数
PTX1.7XLARGE116	腾讯紫霄C100 * 1	16GB * 1	28	11 6	13	230万	28
PTX1.14XLARGE232	腾讯紫霄C100 * 2	16GB * 2	56	23 2	25	470万	48
PTX1.28XLARGE464	腾讯紫霄C100 * 4	16GB * 4	11 2	46 4	50	950万	48
PTX1.56XLARGE928	腾讯紫霄C100 * 8	16GB * 8	22 4	92 8	100	1900万	48

渲染型实例

最近更新时间：2025-04-11 15:45:02

GPU 渲染型实例 适用于 GPU 传统图形图像处理（3D 渲染）应用场景。腾讯云以和 [云服务器 CVM](#) 一致的管理方式，提供快速、稳定、弹性的计算服务。

适用场景

适用于高性能图形处理，3D 渲染。例如：

- 非线性编辑
- 云游戏
- 云手机
- 云桌面
- CloudXR
- 图形图像处理等

渲染型实例总览

GPU 云服务器渲染型系列提供以下实例：

❗ 说明：

下列 NVIDIA GPU 类型实例中，若公共镜像命名中包含“GRID”字样为预装 GRID 驱动的镜像，可应用于渲染场景。

实例	GPU 类型	可用镜像	售卖地域
GA3	GA01	Windows Server 2019 数据中心版 64 位	广州
GNV4v	NVIDIA A10	● Windows Server 2019 数据中心版 64位 中文版 GRID16.2 ● Windows Server 2022 数据中心版 64位 中文版 GRID16.2	广州、上海、北京

GNV4	NVIDIA A10	- CentOS 7.2 及以上 - Ubuntu 16.04 – 20.04 - Windows Server 2019 数据中心版 64位 中文版 GRID16.2 - Windows Server 2022 数据中心版 64位 中文版 GRID16.2	广州、上海、北京、南京、重庆
GN7vw	NVIDIA T4	- TencentOS Server 3.1 (TK4) GRID16.2 - Ubuntu Server 20.04 LTS 64位 GRID16.2 - Ubuntu Server 22.04 LTS 64位 GRID16.2 - Windows Server 2019 数据中心版 64位 中文版 GRID16.2 - Windows Server 2022 数据中心版 64位 中文版 GRID16.2	广州、上海、南京、北京、成都、重庆、中国香港、新加坡、曼谷、法兰克福、首尔、东京、硅谷、弗吉尼亚
GI1	Intel SG1	- CentOS 7.6 64位 + SG1-pv1.3 - CentOS 7.6 64位 + SG1-pv1.4 - CentOS 7.6 64位 + SG1-pv1.5 - CentOS 7.6 64位 + SG1-pv1.6	广州、南京、重庆

渲染型实例选型推荐

腾讯云提供了类型丰富的 GPU 计算实例，可满足不同业务应用场景的需求。请参见下表，并结合实际需求选择合适的计算实例。

GPU 云服务器渲染型实例选型推荐如下表，其中 ✓ 为支持，★ 为推荐。

功能\实例	GA3	GNV4v	GNV4	GN7vw	GI1
图形图像处理	★	★	★	★	★
视频编解码	★	★	★	★	★
深度学习训练	—	—	✓	—	—
深度学习推理	—	—	✓	—	—
科学计算	—	—	—	—	—

注意：

- 以上推荐用途仅供参考，请根据实际需要进行选择。

- NVIDIA 系列 GPU 实例如用作 3D 图形渲染任务（高性能图形处理，视频编解码等），则需安装 GRID Driver 和配置 License Server，安装方法请参见 [安装 NVIDIA GRID 驱动](#)。GNV4v、GNV4、GN7vw 实例可选择已预装 GRID 驱动的指定镜像，无需单独安装 GRID Driver 和配置 License Server。
- GNV4v、GNV4、GN7vw 实例簇提供支持 vDWS/vWS 类型的 vGPU 实例类型，支持 DirectX 和 OpenGL 等图形 API。

实例规格

渲染型 GA3

GA3 适用于图形图像处理（3D 渲染，视频编解码）场景。

⚠ 注意：

- 该实例暂时处于白名单开放状态，请您通过 [售前在线咨询](#) 进行实例购买权限开通。
- 该实例必须使用预装 GA01 驱动镜像方可正常使用。在购买页单击[镜像市场](#)，搜索“预装GA01 GPU 驱动”即可使用。详情可参见 [使用预装 GPU 驱动的镜像](#)。

硬件规格

- **GPU：** 腾讯云星星海 GA01。
- **CPU：** 2.55GHz AMD EPYC™ Milan 处理器，睿频3.5GHz。
- **内存：** 搭配八通道 DDR4 内存。
- **存储：** 可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络：** 默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
GA3.LARGE10	GA01 * 1/6	32GB * 1/6	4	10	1.5	15万	4
GA3.2XLARGE14	GA01 * 1/4	32GB * 1/4	8	14	2	35万	8
GA3.4XLARGE30	GA01 * 1/2	32GB * 1/2	16	30	4	75万	16
GA3.8XLARGE	GA01 * 1	32GB * 1	32	62	7	150万	32

RGE62							
-------	--	--	--	--	--	--	--

渲染型 GNV4v

NVIDIA 实例 GNV4v 配置了 vDWS License 服务器并安装 GRID driver 的渲染型实例，适用于图形图像处理（3D 渲染，视频编解码）场景。使用该实例，您可免除手动配置 GPU 图形图像处理基础环境。

 **注意：**
该实例处于白名单开放状态，请您通过 [售前在线咨询](#) 进行实例购买权限开通。

硬件规格

- **GPU：** NVIDIA[®] A10（FP32 31.2 TFLOPS，INT8 250 TOPS）。
- **CPU：** 2.55GHz AMD EPYC[™] Milan 处理器，睿频3.5GHz。
- **内存：** 搭配八通道 DDR4 内存。
- **存储：** 可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络：** 默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
GNV4v.LARGE24	NVIDIA A10 * 1/4	24GB * 1/4	6	24	3	50万	6
GNV4v.3XLARGE58	NVIDIA A10 * 1/2	24GB * 1/2	14	58	7	110万	14
GNV4v.7XLARGE116	NVIDIA A10 * 1	24GB * 1	28	116	13	230万	28

渲染型 GNV4

NVIDIA 实例 GNV4 配置了 vDWS License 服务器并安装 GRID driver 的渲染型实例（含 GRID 的公共镜像，例如 Windows Server 2022 数据中心版 64位 中文版 GRID16.2），适用于图形图像处理（3D 渲染，视频编解码）场景。使用该实例，您可免除手动配置 GPU 图形图像处理基础环境。

硬件规格

- **GPU：** NVIDIA[®] A10（31.2 TFLOPS 单精度浮点计算，250 INT8 TOPS，500 INT4 TOPS）。
- **CPU：** Intel[®] Xeon[®] Cooper Lake 处理器，基频 3.4 GHz，睿频3.8 GHz。

- **内存：**搭配六通道 DDR4 内存。
- **存储：**可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络：**默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
GNV4.3XLA RGE44	NVIDIA A10 * 1	24GB	12	44	2	50万	12
GNV4.6XLA RGE88	NVIDIA A10 * 2	24GB * 2	24	88	4	100万	24
GNV4.12XLA RGE176	NVIDIA A10 * 4	24GB * 4	48	176	7	210万	32
GNV4.24XLA RGE352	NVIDIA A10 * 8	24GB * 8	96	352	13	420万	32
GNV4.48XLA RGE704	NVIDIA A10 * 16	24GB * 16	192	704	25	850万	32

渲染型 GN7vw

NVIDIA 实例 GN7vw 是在 GN7 基础上配置 vDWS License 服务器并安装 GRID driver 的渲染型实例，适用于图形图像处理（3D 渲染，视频编解码）场景。使用该实例，您可免除手动配置 GPU 图形图像处理基础环境。

 **注意：**
该实例处于白名单开放状态，请您通过 [售前在线咨询](#) 进行实例购买权限开通。

硬件规格

- **GPU：**NVIDIA® Tesla® T4（FP32 8.1 TFLOPS, 130 INT8 TOPS, 260 INT4 TOPS）。
- **CPU：**Intel® Xeon® Platinum 8255C CPU，主频 2.5 GHz。
- **内存：**DDR4，内存速率达2666MT/s。
- **存储：**可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络：**默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽	网络收发包 (PPS)	队列数
----	-----	--------	------	----------	------	-------------	-----

					(Gbps)		
GN7vw.LARGE16	NVIDIA T4 * 1/4	16GB * 1/4	4	16	2	50万	8
GN7vw.2XLA RGE32	NVIDIA T4 * 1/2	16GB * 1/2	8	32	4	80万	8
GN7vw.4XLA RGE64	NVIDIA T4 * 1	16GB * 1	16	64	7	150万	8

渲染型 G11

GPU 渲染型 G11 实例搭载 H3C XG310 加速卡，单张加速卡包含 4 颗 Intel SG1 芯片，适用于 Android 云游戏、Android 云应用、视频转码处理等场景。

⚠ 注意：

该实例处于白名单开放状态，请您通过 [售前在线咨询](#) 进行实例购买权限开通。

适用场景

- Android 云手机
- Android 云游戏
- Android 云应用
- 视频转码

硬件规格

- **GPU：**Intel® SG1，采用 H3C XG310 加速卡，单张加速卡包含4颗 SG1 芯片。
- **CPU：**Intel® Xeon® Platinum 8255c CPU，主频2.5GHz。
- **存储：**可选择 [云硬盘类型](#)，如需 [扩容](#) 可新建弹性云盘进行挂载。
- **网络：**默认网络优化，实例网络性能与规格对应。[公网网络](#) 可按需配置。

规格	GPU	GPU 显存	vCPU	内存 (GiB)	内网带宽 (Gbps)	网络收发包 (PPS)	队列数
G11.10XLARGE160	H3C XG310 * 1 (Intel SG1 芯片 * 4)	32GB * 1 (8GB * 4)	42	160	13	250万	32

GI1.21XLAR GE320	H3C XG310 * 2 (Intel SG1 芯 片 * 8)	32GB * 2 (8GB * 8)	84	320	25	600万	32
---------------------	--	-----------------------------	----	-----	----	------	----