高性能计算集群

实例规格

腾讯云

版权所有：腾讯云计算（北京）有限责任公司

【版权声明】

©2013-2024 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】

腾讯云

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。

实例规格

最近更新时间：2024-04-25 16:13:01

高性能计算集群以高性能云服务器为节点，通过 RDMA（Remote Direct Memory Access）互联，提供了高带宽和极低延迟的网络服务，大幅提升网络性能满足大规模高性能计算、人工智能、大数据推荐等应用的并行计算需求。

实例总览

高性能计算集群型可提供以下实例规格：

售卖情况	实例	实例类型	GPU	可用镜像
主售	HCCPNV5	GPU型	Nvidia H800	• TencentOS Server 3.1 (TK4) UEFI版
	HCCPNV5v	GPU型	Nvidia H800	• TencentOS Server 2.4 (TK4)
	HCCPNV4sne	GPU型	Nvidia A800	• TencentOS Server 2.4 (TK4)
	HCCPNV4sn	GPU型	Nvidia A800	• TencentOS Server 2.4 (TK4)
	HCCPNV4h	GPU型	Nvidia A100	• TencentOS Server 2.4 (TK4) • Ubuntu Server 18.04 LTS • CentOS 7.6
在售	HCCG5vm	GPU型	Nvidia V100	• TencentOS Server 2.4 (TK4) • Ubuntu Server 18.04 LTS • CentOS 7.6

版权所有：腾讯云计算（北京）有限责任公司

	HCCG5v	GPU型	Nvidia V100	• TencentOS Server 2.4 (TK4) • Ubuntu Server 18.04 LTS • CentOS 7.6
	HCCS5	标准型	–	• TencentOS Server 2.4 (TK4) • Ubuntu Server 18.04 LTS • CentOS 7.6
	HCCIC5	计算型	–	• TencentOS Server 2.4 (TK4) • Ubuntu Server 18.04 LTS • CentOS 7.6

实例规格

请参照以下介绍选择符合您业务需求的实例规格，确保所选规格满足您业务的最低CPU、内存、GPU等需求。

GPU 型 HCCPNV5

GPU 型 HCCPNV5 实例是最新一代的 GPU 型高性能计算实例，搭载 NVIDIA® H800 Tensor Core GPU。GPU 卡间支持400GB/s NVLink 互联，实例间支持 3.2Tbps RDMA 星脉网络互联，性能强大。

> **说明：**
> 该实例暂时处于白名单开放，请联系您的售前经理进行实例购买权限开通。

使用场景

HCCPNV5 具有强大的浮点运算能力 ，适用于大规模 AI 和科学计算场景：

- 大规模深度学习训练、大数据推荐等。
- HPC 应用，例如计算金融、材料量子模拟、分子建模等。

硬件规格

- **CPU**：2.6GHz Intel® Xeon® Sapphire Rapids 处理器，睿频3.1GHz。
- **GPU**：搭载 8 × NVIDIA® H800 NVLink® 80GB（FP32 64 TFLOPS, TF32 494 TFLOPS, BF16 989 TFLOPS, 400GB/s NVLink®）。

- **内存**：搭配八通道 DDR5 内存。
- **存储**：8 × 6400 GB NVMe SSD 高性能本地存储。支持 云硬盘 作为系统盘和数据盘，可按需 扩容 。
- **网络**：支持 100Gbps 内网带宽，支持 3.2Tbps 的低延迟 RDMA 网络，专用于高性能计算集群内通信，具有超高网络收发包能力。公网网络 可按需配置。支持弹性网卡挂载。

规格	vCPU	内存(GiB)	主频/睿频(GHz)	GPU	GPU显存	RDMA配置	内网带宽能力(Gbps)	网络收发包(pps)	队列数	连接数	本地存储
HCCPNV5	192	2048	2.6/3.1	Nvidia H800×8	80GB×8	3.2Tbps RoCEv2	100	4500万	32	1600万	8 × 6400 GB NVMe SSD

> **说明**
> GPU 驱动：NVIDIA H800 系列建议安装 NVIDIA Tesla 535 版本以上驱动，推荐您安装 535.54.03(Linux)/536.25(Windows) 版本驱动，驱动版本信息请参见 NVIDIA 官方文档 。

GPU 型 HCCPNV5v

GPU 型 HCCPNV5v 实例是最新一代的 GPU 型高性能计算实例，搭载 NVIDIA® H800 Tensor Core GPU。GPU 卡间支持400GB/s NVLink 互联，实例间支持 3.2Tbps RDMA 星脉网络互联，性能强大。

> **说明：**
> 该实例暂时处于白名单开放，请联系您的售前经理进行实例购买权限开通。

使用场景

HCCPNV5v 具有强大的浮点运算能力 ，适用于大规模 AI 和科学计算场景：

- 大规模深度学习训练、大数据推荐等。
- HPC 应用，例如计算金融、材料量子模拟、分子建模等。

硬件规格

- **CPU**：2.6GHz Intel® Xeon® Sapphire Rapids 处理器，睿频3.1GHz。
- **GPU**：搭载 8 × NVIDIA® H800 NVLink® 80GB（FP32 64 TFLOPS, TF32 494 TFLOPS, BF16 989 TFLOPS, 400GB/s NVLink®）。
- **内存**：搭配八通道 DDR5 内存。
- **存储**：8 × 6400 GB NVMe SSD 高性能本地存储。支持 云硬盘 作为系统盘和数据盘，可按需 扩容 。
- **网络**：支持 100Gbps 内网带宽，支持 3.2Tbps 的低延迟 RDMA 网络，专用于高性能计算集群内通信，具有超高网络收发包能力。公网网络 可按需配置。支持弹性网卡挂载。

规格	vCPU	内存(GiB)	主频/睿频(GHz)	GPU	GPU显存	RDMA配置	内网带宽能力(Gbps)	网络收发包(pps)	队列数	连接数	本地存储
HCCPNV5v	172	1939	2.6/3.1	Nvidia H800 × 8	80GB × 8	3.2Tbps RoCEv2	100	1500万	48	1600万	8 × 6400 GB NVMe SSD

> **说明**
> GPU 驱动：NVIDIA H800 系列建议安装 NVIDIA Tesla 535 版本以上驱动，推荐您安装 535.54.03(Linux)/536.25(Windows) 版本驱动，驱动版本信息请参见 NVIDIA 官方文档 。

GPU 型 HCCPNV4sne

GPU 型 HCCPNV4sne 实例是次新一代的 GPU 型高性能计算实例，搭载 NVIDIA® A800 Tensor Core GPU。GPU 卡间支持400GB/s NVLink 互联，实例间支持 1.6Tbps RDMA 星脉网络互联，性能强大。

> **说明：**
> 该实例暂时处于白名单开放，请联系您的售前经理进行实例购买权限开通。

使用场景

HCCPNV4sne 具有强大的浮点运算能力 ，适用于大规模 AI 和科学计算场景：

- 大规模深度学习训练、大数据推荐等。
- HPC 应用，例如计算金融、材料量子模拟、分子建模、基因测序等。

硬件规格

- **CPU**：2.7GHz Intel® Xeon® Ice Lake 处理器，睿频3.3GHz。
- **GPU**：搭载 8 × NVIDIA® A800 NVLink® 80GB（FP64 9.7 TFLOPS, TF32 156 TFLOPS, BF16 312 TFLOPS, 400GB/s NVLink®）。
- **内存**：搭配八通道 DDR4 内存。
- **存储**：4 × 6400 GB NVMe SSD 高性能本地存储。支持 云硬盘 作为系统盘和数据盘，可按需 扩容 。
- **网络**：支持 100Gbps 内网带宽，支持 1.6Tbps 的低延迟 RDMA 网络，专用于高性能计算集群内通信，具有超高网络收发包能力。公网网络 可按需配置。支持弹性网卡挂载。

规格	vCPU	内存(GiB)	主频/睿频(GHz)	GPU	GPU显存	RDMA配置	内网带宽能力(Gbps)	网络收发包(pps)	队列数	连接数	本地存储
HCCPNV4sne	124	1929	2.7/3.3	Nvidia A800 × 8	80GB × 8	1.6Tbps RoCEv2	100	1500万	48	1600万	4 × 6400 GB NVMe SSD

> **说明：**
> GPU 驱动：NVIDIA A800 系列需要安装 NVIDIA Tesla 450 版本以上驱动，推荐您安装 460.32.03(Linux)/461.33(Windows) 版本驱动，驱动版本信息请参见 NVIDIA 官方文档 。

GPU 型 HCCPNV4sn

GPU 型 HCCPNV4sn 实例是次新一代的 GPU 型高性能计算实例，搭载 NVIDIA® A800 Tensor Core GPU。GPU 卡间支持400GB/s NVLink 互联，实例间支持 800Gbps RDMA 星脉网络互联，性能强大。

> **说明：**
> 该实例暂时处于白名单开放，请联系您的售前经理进行实例购买权限开通。

使用场景

HCCPNV4sn 具有强大的浮点运算能力 ，适用于大规模 AI 和科学计算场景：

- 大规模深度学习训练、大数据推荐等。
- HPC 应用，例如计算金融、材料量子模拟、分子建模、基因测序等。

硬件规格

- **CPU**：2.55GHz AMD EPYC™ Milan 处理器，睿频3.5GHz。
- **GPU**：搭载 8 × NVIDIA® A800 NVLink® 80GB（FP64 9.7 TFLOPS, TF32 156 TFLOPS, BF16 312 TFLOPS, 400GB/s NVLink®）。
- **内存**：搭配八通道 DDR4 内存。
- **存储**：2 × 7680 GB NVMe SSD 高性能本地存储。支持 云硬盘 作为系统盘和数据盘，可按需 扩容 。
- **网络**：支持 100Gbps 内网带宽，支持 800Gbps 的低延迟 RDMA 网络，专用于高性能计算集群内通信，具有超高网络收发包能力。公网网络 可按需配置。支持弹性网卡挂载。

规格	vCPU	内存(GiB)	主频/睿频(GHz)	GPU	GPU显存	RDMA配置	内网带宽能力(Gbps)	网络收发包(pps)	队列数	连接数	本地存储
HCCPNV4sn	232	1929	2.55/3.5	Nvidia A800 × 8	80GB × 8	800Gbps RoCEv2	100	1900万	48	1600万	2 × 7680 GB NVMe SSD

> **说明：**
> GPU 驱动：NVIDIA A800 系列需要安装 NVIDIA Tesla 450 版本以上驱动，推荐您安装 460.32.03(Linux)/461.33(Windows) 版本驱动，驱动版本信息请参见 NVIDIA 官方文档 。

GPU 型 HCCPNV4h

GPU 型 HCCPNV4h 实例是次新一代的 GPU 型高性能计算实例，搭载 NVIDIA® A100 Tensor Core GPU，并基于 NVMe SSD 实例存储，提供低延迟、超高 IOPS、高吞吐量的存储资源，性能强大。

使用场景

HCCPNV4h 具有强大的双精度浮点运算能力 ，适用于大规模 AI 和科学计算场景：

- 大规模机器学习训练、大数据推荐。
- HPC 应用，例如计算金融、材料量子模拟、分子建模、基因测序等。

硬件规格

- **CPU**：2.6GHz AMD EPYC™ ROME 处理器，睿频3.3GHz。
- **GPU**：搭载 8 × NVIDIA® A100 NVLink® 40GB（FP64 9.7 TFLOPS, TF32 156 TFLOPS, BF16 312 TFLOPS, 600GB/s NVLink®）。
- **内存**：搭配八通道 DDR4 内存。
- **存储**：配备 1 × 480 GB SATA SSD 本地系统盘及 4 × 3200 GB NVMe SSD 高性能本地存储。不支持云硬盘挂载。
- **网络**：支持25Gbps内网带宽，支持 100Gbps 的低延迟 RDMA 网络，专用于高性能计算集群内通信，具有超高网络收发包能力。公网网络 可按需配置。不支持弹性网卡挂载。

规格	vCPU	内存(GiB)	主频/睿频(GHz)	GPU	GPU显存	RDMA配置	内网带宽能力(Gbps)	网络收发包(pps)	队列数	连接数	本地存储
HCCPNV4h	192	1024	2.6/3.3	Nvidia A100	40GB × 8	100Gbps RoCEv2	25	1000万	16	200万	1 × 480 GB SATA SSD 及 4 × 3200 GB NVMe SSD

				× 8							

> **说明：**
> GPU 驱动：NVIDIA A100 系列需要安装 NVIDIA Tesla 450 版本以上驱动，推荐您安装 460.32.03(Linux)/461.33(Windows) 版本驱动，驱动版本信息请参见 NVIDIA 官方文档 。

GPU 型 HCCG5vm

GPU 型 HCCG5vm 实例搭载 NVIDIA® Tesla® V100 GPU，并基于 NVMe SSD 实例存储，提供低延迟、超高 IOPS、高吞吐量的存储资源，性能强大。

使用场景

- 大规模机器学习训练、大数据推荐。
- HPC 应用，例如计算金融、材料量子模拟、分子建模、基因测序等。

硬件规格

- **CPU**：2.5GHz Intel® Xeon® Cascade Lake 处理器，睿频3.1GHz。
- **GPU**：搭载 8 × NVIDIA® Tesla® V100 GPU（FP64 7.8 TFLOPS, FP32 15.7 TFLOPS，300GB/s NVLink®）。
- **内存**：搭配六通道 DDR4 内存。
- **存储**：配备 1 × 480 GB SATA SSD 本地系统盘及 4 × 3200 GB NVMe SSD 高性能本地存储。不支持云硬盘挂载。
- **网络**：支持25Gbps内网带宽，支持100Gbps的低延迟 RDMA 网络，专用于高性能计算集群内通信，具有超高网络收发包能力。公网网络 可按需配置。不支持弹性网卡挂载。

规格	vCPU	内存(GiB)	主频/睿频(GHz)	GPU	GPU显存	RDMA配置	内网带宽能力(Gbps)	网络收发包(pps)	队列数	连接数	本地存储
HCCG5vm	96	768	2.5/3.1	Nvidia V10	32GB × 8	100Gbps Ro	25	1000万	16	200万	1 × 480 GB SATA SSD 及 4 × 3200 GB

				0 × 8		CE v2					NVMe SSD

GPU 型 HCCG5v

GPU 型 HCCG5v 实例搭载 NVIDIA® Tesla® V100 GPU，并基于 NVMe SSD 实例存储，提供低延迟、超高 IOPS、高吞吐量的存储资源，性能强大。

使用场景

- 大规模机器学习训练、大数据推荐。
- HPC 应用，例如计算金融、材料量子模拟、分子建模、基因测序等。

硬件规格

- **CPU**：2.5GHz Intel® Xeon® Cascade Lake 处理器，睿频3.1GHz。
- **GPU**：搭载 8 × NVIDIA® Tesla® V100 GPU（FP64 7.8 TFLOPS, FP32 15.7 TFLOPS，300GB/s NVLink®）。
- **内存**：搭配六通道 DDR4 内存。
- **存储**：配备 1 × 480 GB SATA SSD 本地系统盘及 4 × 3200 GB NVMe SSD 高性能本地存储。不支持云硬盘挂载。
- **网络**：支持25Gbps内网带宽，支持100Gbps的低延迟 RDMA 网络，专用于高性能计算集群内通信，具有超高网络收发包能力。公网网络 可按需配置。不支持弹性网卡挂载。

规格	vCPU	内存(GiB)	主频/睿频(GHz)	GPU	GPU显存	RDMA配置	内网带宽能力(Gbps)	网络收发包(pps)	队列数	连接数	本地存储
HCCG5v	96	384	2.5/3.1	Nvidia V100 × 8	32 GB × 8	100 Gbps RoCE v2	25	1000万	16	200万	1 × 480 GB SATA SSD 及 4 × 3200 GB NVMe SSD

标准型 HCCS5

标准型 HCCS5 实例搭载2.5GHz基准主频 CPU，适合通用多核批处理、多核高性能计算应用等计算密集型应用。

使用场景

- 大规模高性能计算应用。
- HPC 应用，例如流体力学分析、工业仿真、分子建模、基因测序、气象分析等。

硬件规格

- **CPU**：2.5GHz Intel® Xeon® Cascade Lake 处理器，睿频3.1GHz。
- **内存**：搭配六通道 DDR4 内存。
- **存储**：配备 1 × 480 GB SATA SSD。不支持云硬盘挂载。
- **网络**：支持25Gbps内网带宽，支持100Gbps的低延迟 RDMA 网络，专用于高性能计算集群内通信，具有超高网络收发包能力。公网网络 可按需配置。不支持弹性网卡挂载。

规格	vCPU	内存(GiB)	主频/睿频(GHz)	RDMA 配置	内网带宽能力(Gbps)	网络收发包(pps)	队列数	连接数	本地存储
HCCS5	96	384	2.5/3.1	100 Gbps RoCEv2	25	1000万	16	200万	1 × 480 GB SATA SSD

计算型 HCCIC5

高 IO 计算型 HCCIC5 实例搭载3.2GHz基准主频 CPU，具有高单核计算性能，并基于 NVMe SSD 实例存储，提供低延迟、超高 IOPS 的存储资源。适合批处理、流体力学、结构仿真等计算密集型及IO密集型应用。

使用场景

- 大规模高性能计算应用。
- HPC 应用，例如流体力学分析、工业仿真、分子建模、基因测序、气象分析等。

硬件规格

- **CPU**：3.2GHz Intel® Xeon® Cascade Lake 处理器，睿频3.7GHz。
- **内存**：搭配六通道 DDR4 内存。
- **存储**：配备 2 × 480 GB SATA SSD (RAID1) 本地系统盘及 2 × 3840 GB NVMe SSD 高性能本地存储。不支持云硬盘挂载。

版权所有：腾讯云计算（北京）有限责任公司

- **网络：支持25Gbps内网带宽，支持100Gbps的低延迟 RDMA 网络，专用于高性能计算集群内通信，具有超高网络收发包能力。公网网络 可按需配置。不支持弹性网卡挂载。**

规格	vCPU	内存(GiB)	主频/睿频(GHz)	RDMA 配置	内网带宽能力(Gbps)	网络收发包(pps)	队列数	连接数	本地存储
HCCIC5	64	384	3.2/3.7	100 Gbps RoCEv2	25	1000万	16	200万	2 × 480 GB SATA SSD (RAID1)及 2 × 3840 GB NVMe SSD