Alibaba Cloud | WORKSPACE

Alibaba Cloud Workspace NVIDIA RTX 5880-powered Series Workstations –Introduction and Sales Guide

Alibaba Cloud | WORKSPACE

Alibaba Cloud Workspace NVIDIA RTX 5880-powered Series Workstations: Optimal Solution for Graphics Workstation Cloud Computers

Powered by **NVIDIA enterprise-level GPUs dedicated to Alibaba Cloud Workspace**, the 5880-powered series workstations provide **Graphics Workstation** cloud computers designed to meet the needs of graphic design, image rendering, simulation, and inference scenarios. NVIDIA RTX 5880 has higher performance and lower unit costs than other GPUs with the same memory size. 5880-powered series workstations are available in Pay-as-you-go, Monthly Subscription - Unlimited, and Subscription Hourly Plan modes.

Capability Improvement

Robust performance

- Higher GPU performance than GPUs with the same memory size
- Superior stability and reliability enabled by high software compatibility

Graphics capabilities

- Enterprise-level GPUs
- Up to 3.7 GHz of CPU clock speed
- 48 GiB GPU memory

Architecture upgrade

- Alibaba Cloud Workspace dedicated host
- Resource assurance
- Inventory guarantee

Advantages

Alibaba Cloud Workspace vs. Other DaaS vendors

- Alibaba Cloud Workspace provides industry-leading GPU performance.
- Other cloud vendors still use Tesla T4 GPUs for graphics-based cloud computers.

NVIDIA RTX 5880 vs. NVIDIA L20 and A10

- Benchmark results show that NVIDIA RTX 5880 outperforms other GPUs with the same memory size.
- NVIDIA RTX 5880 provides a smoother experience and higher frame rate.

Scenarios

Software Development, Data Warehouse, Business Intelligence Applications
Applicable GPU memory: 4 GiB.

Complex component design, autonomous driving development, lightweight simulation, and game development.
Applicable GPU memory: 8 GiB and 16 GiB.

Modeling design of large industrial products and heavy simulation.
Applicable GPU memory: 24 GiB and higher.

Embodied AI, data labeling, simulation, and inference.
Applicable GPU memory: 48 GiB and higher.

CONTENT

Contents

Introduction to 5880-powered Series Workstations

Powered by NVIDIA enterprise-level GPUs dedicated to Alibaba Cloud Workspace, the 5880-powered series workstations provide cloud computers designed to meet the needs of graphic design, image rendering, simulation, and inference scenarios.

CPU: AMD-Genoa 9T24 (up to 3.7 GHz)

GPU: NVIDIA RTX 5880

Benefits

Diverse specifications

- **Specifications** with 4 GiB to 192 GiB GPU memory
- Flexible configuration changes
- Various GPU utilization scenarios

Robust performance

- Enterprise-level GPUs
- Higher GPU performance than GPUs with the same memory size
- Superior stability and reliability enabled by high software compatibility

Cost-effectiveness

- **Lower unit prices compared with GPUs with the same memory size**
- Flexible purchase options: Pay-as-you-go, Subscription Hourly Plan, and Monthly Subscription - Unlimited
- Subscription Hourly Plan: monthly consumption capped at full subscription price

excellent user experience

- **Nearby access in 8 regions worldwide**
- Data security ensured by the security protection capabilities of cloud computers

Inventory guarantee

- **Exclusive inventory** to prevent resource conflicts with other product lines
- Flexible GPU provisioning

12 specifications are available to support different graphics rendering scenarios.

vCPU	Memory	GPU (GiB)	Applicable scenario	Scenario characteristic
8	16	4	Software Development, Data Warehouse, Business Intelligence Applications	Moderate VRAM requirements, not intended for professional design scenarios
16	32	8	Complex component design, autonomous driving development, lightweight simulation, and game development	GPU memory requirements are high, as they directly determine the drawing size that can be opened.
32	64	16		
32	64	24	Modeling design of large industrial products, heavy simulation, and robot development	High levels of GPU memory, CPU performance, and system memory are required.
64	128	24		
32	64	48	Large workstations and AI inference	Rendering, design, and simulation of complex ultra-large models and scientific computing are required.
64	128	48		
64	256	48		
64	256	96		
96	384	192		

RTX Ada 5880: The Ideal Graphics Workstation GPUs for Embodied AI

Isaac Sim Requirements

Hint

Before installing Isaac Sim, install and run the lightweight app Isaac Sim Compatibility Checker to check if the machine meets the system requirements and compatibility.

System Requirements

Element	Minimum Spec	Good	Ideal
OS	Ubuntu 20.04/22.04 Windows 10/11	Ubuntu 20.04/22.04 Windows 10/11	Ubuntu 20.04/22.04 Windows 10/11
CPU	Intel Core i7 (7th Generation) AMD Ryzen 5	Intel Core i7 (9th Generation) AMD Ryzen 7	Intel Core i9, X-series or higher AMD Ryzen 9, Threadripper or higher
Cores	4	8	16
RAM	32GB [1]	64GB [1]	64GB [1]
Storage	50GB SSD	500GB SSD	1TB NVMe SSD
GPU	GeForce RTX 3070	GeForce RTX 4080	RTX Ada 6000
VRAM	8GB [1] [2]	16GB	48GB

GPU specification	NVIDIA RTX 5880	NVIDIA RTX 6000	GeForce RTX 4080	NVIDIA L20
GPU memory	48 GiB GDDR6	48 GiB GDDR6	16 GiB GDDR6	48 GiB GDDR6
GPU memory bus width	384 bits	384 bits	256 bits	384 bits
GPU memory bandwidth	960 Gbit/s	960 Gbit/s	768 Gbit/s	864 Gbit/s
Error correction code (ECC)	Yes	Yes	No	Yes
CUDA cores	14,080	18,176	9,728	11,776
NVIDIA Fourth-generation Tensor Cores	440	568	780	576
NVIDIA Third-generation RT Cores	110	142	113	59.5
Single-precision performance	69.3 TFLOPS	91.1 TFLOPS	45	59.5
RT Core performance	160.2 TFLOPS	210.6 TFLOPS	113	144
Tensor performance	1,108.4 TFLOPS	1,457.0 TFLOPS	578.5	714

RTX Ada 5880 GPUs provide performance comparable to RTX Ada 6000 GPUs and are available in China.

The Third-generation RT Cores accelerate ray tracing, enabling real-time rendering in NVIDIA Omniverse with a **50% lower latency** compared with similar-level platforms.

Comparison of Theoretical Performance Among NVIDIA RTX 5880, L20, and A10

Compared with NVIDIA L20 and A10, NVIDIA RTX 5880 **leads in every aspect** and is more suitable for graphics rendering, industrial software, and architectural visualization.

GPU specification	NVIDIA RTX 5880	NVIDIA L20	NVIDIA A10
Architecture	Ada Lovelace	Ada Lovelace	NVIDIA Ampere
Core purpose	Professional graphics rendering, design, and AI	AI computing and data center acceleration	Interactive rendering and virtual workstations
GPU memory	48 GiB GDDR6	48 GiB GDDR6	24 GiB GDDR6
GPU memory bandwidth	**960 Gbit/s**	864 Gbit/s	600 Gbit/s
CUDA cores	**14,080**	11,776	9,216
FP32 • The **gold standard** for graphics rendering, especially in high-fidelity scenarios, such as 3A games and cinematic computer graphics (CG). • Use scenarios: graphics rendering, traditional deep learning inference, and scientific computing. • Benefits: high precision and versatility.	**69.3**	59.5	31.2
TF32	**138.5**	59.5	62.5
FP16	**277**	119	125
FP8	**554**	239	-

1. Comparison of NVIDIA RTX 5880, L20, and A10 using SPECviewperf

GPU memory size	Software	NVIDIA RTX 5880	NVIDIA L20	NVIDIA A10
24 GiB GPU memory	sw-07	**291.67**	272.39	285.83
	maya-06	**423.07**	365.26	367.82
	snx-04	**541.53**	454.76	457.74
	catia-06	**115.39**	101.06	102.65
48 GiB GPU memory	sw-07	**563.73**	480.25	-
	maya-06	**643.02**	627.30	-
	snx-04	**961.49**	808.46	-
	catia-06	**175.28**	135.93	-

2. Comparison of NVIDIA RTX 5880, L20, and A10 using 3DMark

GPU memory size	NVIDIA RTX 5880	NVIDIA L20	NVIDIA A10
24 GiB GPU memory	**32,018**	25,393	28,791
48 GiB GPU memory	**50,829**	43,800	-

3. Comparison of NVIDIA RTX 5880, L20, and A10 using V-ray

GPU memory size	NVIDIA RTX 5880	NVIDA L20	NVIDIA A10
24 GiB GPU memory	**2,471**	2,263	2,042
48 GiB GPU memory	**5,381**	4,934	-

4. Comparison of NVIDIA RTX 5880, L20, and A10 using Unigine Heaven

GPU memory size	NVIDIA RTX 5880	NVIDIA L20	NVIDIA A10
24 GiB GPU memory	**8,161**	7,714	8,097
48 GiB GPU memory	**8,485**	7,700	-

Performance Comparison Between NVIDIA RTX 5880 and L20 on AI Models

The overall performance of 5880-powered professional workstations (48 GiB GPU memory) is approximately 20% higher than that of L20-powered workstations running on Qwen2.5-7B and Llama3-8B.

Model	Precision	batch_size	input_token_length	output_token_length	NVIDIA RTX 5880			NVIDIA L20		
					Throughput (tokens/s)	latency_first_token (ms/token)	throughput_rest_tokens (tokens/s)	Throughput (tokens/s)	latency_first_token (ms/token)	throughput_rest_tokens (tokens/s)
Qwen2.5-7B	Float16	1	512	512	58.82	98.30	59.38	48.02	95.25	48.36
		2	512	512	115.07	171.04	117.10	91.18	158.87	92.31
		4	512	512	222.80	327.76	230.59	177.41	298.27	181.76
		8	512	512	420.13	643.82	448.96	334.79	571.12	350.50
		16	512	512	757.09	1,274.95	856.53	621.99	1,119.92	678.47
		32	512	512	1,204.32	2,549.88	1,479.23	978.27	2,219.86	1,125.54
		64	512	512	1,835.65	5,112.52	2,567.37	1,556.74	4,417.30	1,966.35
llama3-8B	Float16	1	512	512	39.24	98.71	39.55	32.07	99.26	32.26
		2	512	512	107.69	185.7	109.62	85.48	173.94	86.57
		4	512	512	207.92	352.89	215.23	165.73	310.84	169.67
		8	512	512	389.03	670.61	414.68	313.95	571.34	327.68
		16	512	512	693.67	1,264.02	775.29	569.05	1,082.38	614.11
		32	512	512	1,089.95	2,510.63	1,305.95	865.1	2,135.53	973.14
		64	512	512	1,622.70	5,019.42	2,155.25	1,362.59	4,238.26	1,650.89

Operation Experience Comparison Between NVIDIA RTX 5880 and L20 (48 GiB GPU memory) in Design Software

Conclusion: NVIDIA RTX 5880 has higher performance than NVIDIA L20 in rendering smoothness and actual rendering frame rate (see the status bar for details).

Alibaba Cloud | WORKSPACE

Billing and Pricing

GPU specification	GPU specification	Monthly catalog price (USD per unit)	Monthly price (360 hours)	Monthly price (250 hours)	Hourly catalog price (USD per unit)
Graphics Workspace – 8 vCPUs, 16 GiB Memory, 4 GiB GPU Memory	eds.graphics_flagship_pro.8c16g.4g	160	108	88	0.400489
Graphics Workstation (Ultimate)– 16 vCPUs, 32 GiB Memory, 8 GiB GPU Memory	eds.graphics_flagship_pro.16c32g.8g	297	198	164	0.742594
Graphics Workstation (Ultimate)– 32 vCPUs, 64 GiB Memory, 16 GiB GPU Memory	eds.graphics_flagship_pro.32c64g.16g	551	354	293	1.377932
Graphics Workstation (Ultimate)– 32 vCPUs, 64 GiB Memory, 24 GiB GPU Memory	eds.graphics_flagship_pro.32c64g.24g	786	530	432	1.964398
Graphics Workstation (Ultimate)– 64 vCPUs, 128 GiB Memory, 24 GiB GPU Memory	eds.graphics_flagship_pro.64c128g.24g	981	648	540	2.453120
Professional Workstation– 32 vCPUs, 64 GiB Memory, 48 GiB GPU Memory	eds.graphics_flagship_elite.32c64g.48g1x	1,177	784	647	2.941842
Professional Workstation– 64 vCPUs, 128 GiB Memory, 48 GiB GPU Memory	eds.graphics_flagship_elite.64c128g.48g1x	1,568	1,058	862	3.919286
Professional Workstation– 64 vCPUs, 256 GiB Memory, 48 GiB GPU Memory	eds.graphics_flagship_elite.64c256g.48g1x	1,959	1,332	1,077	4.896729
Professional Workstation– 64 vCPUs, 256 GiB Memory, 96 GiB GPU Memory	eds.graphics_flagship_elite.64c256g.96g2x	2,936	1,957	1,615	7.340338
Professional Workstation– 96 vCPUs, 384 GiB Memory, 192 GiB GPU Memory	eds.graphics_flagship_elite.96c384g.192g4x	5,673	3,717	3,120	14.182444

- Supported regions:

  Regions in the Chinese mainland: China (Beijing), China (Shanghai), China (Hangzhou), China (Shenzhen),, and China (Ulanqab).

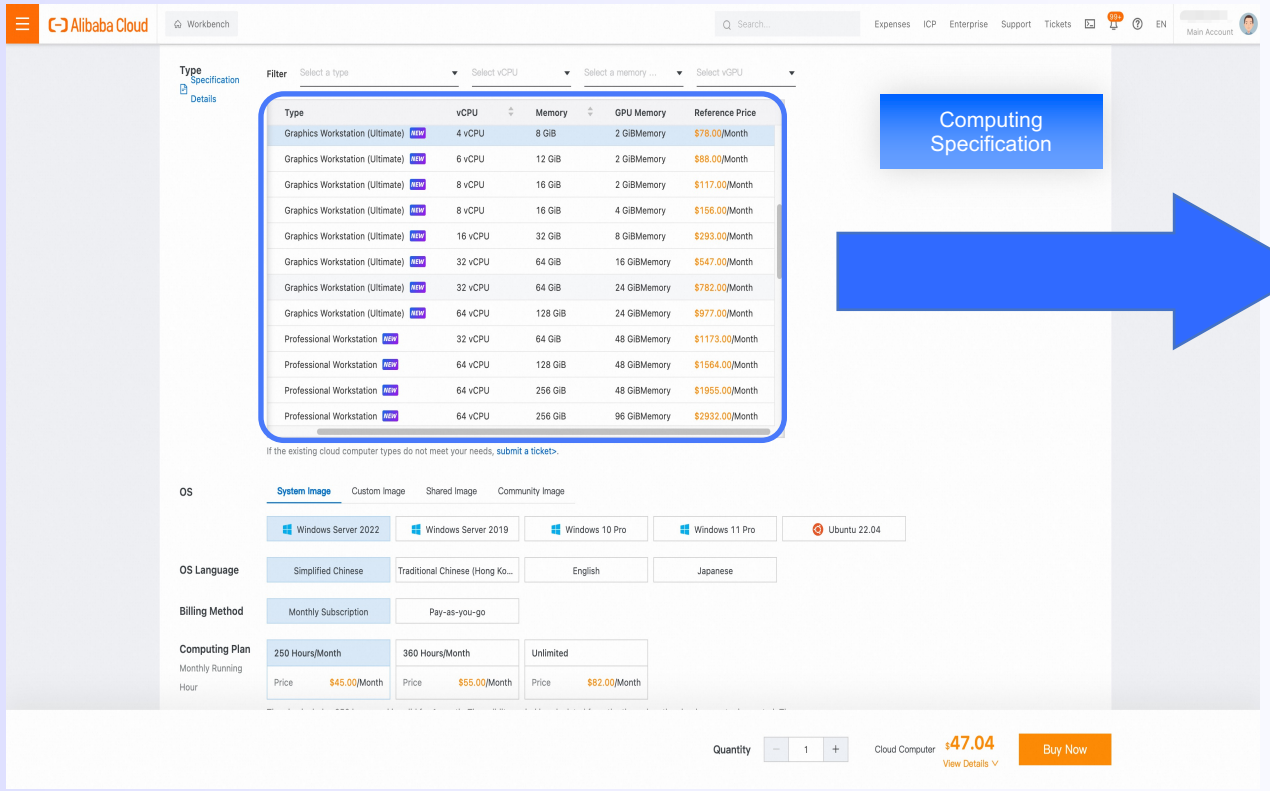
  Regions outside the Chinese mainland: China (Hong Kong), Singapore, and Japan (Tokyo).
- Supported billing methods: Subscription - Unlimited, 360 Hours/Month, 250 Hours/Month, and pay-as-you-go.
- Hibernation is not supported. If you select a subscription hourly plan when purchasing a cloud computer, the automatic shutdown upon disconnection policy is automatically applied. After purchase, you can disable or modify the policy in the Elastic Desktop Service console.
- The catalog price is approximately 30% lower than that of a Graphics Workstation (Ultimate)(NVIDIA L20) cloud computer with the same specifications, and 25% lower than a Graphics Workstation - Basic (NVIDIA T4) or Graphics Workstation - Standard (NVIDIA A10) cloud computer with the same specifications.
- In the previous table, the prices of Windows cloud computers in the China (Hong Kong) and Singapore regions are listed.

Alibaba Cloud | WORKSPACE

Buy Page

Elastic Desktop Service > Custom Purchase

https://eds.console.aliyun.com/edsbuy/createdesktops/cn-hongkong



- Computing specification
- Graphics Workspace– 8 vCPUs, 16 GiB Memory, 4 GiB GPU Memory
- Graphics Workstation (Ultimate)– 16 vCPUs, 32 GiB Memory, 8 GiB GPU Memory
- Graphics Workstation (Ultimate)– 32 vCPUs, 64 GiB Memory, 16 GiB GPU Memory
- Graphics Workstation (Ultimate)– 32 vCPUs, 64 GiB Memory, 24 GiB GPU Memory
- Graphics Workstation (Ultimate)– 64 vCPUs, 128 GiB Memory, 24 GiB GPU Memory

THANK YOU