

# TG668V4

## AI Server for Training and Inference



TG668V4 is versatile 6U 8-card AI server that combines training and inference. Based on the Intel® Xeon® 6 processor, supports variety of mainstream AI accelerator cards in the industry and features excellent performance, flexible architecture, strong scalability, rich configuration, and high reliability.

### New Platform Superior Performance

- Adopts Intel Xeon 6 processor platform, with comprehensive upgrades to CPU computing power and memory bandwidth, fully unleashing GPU computing power;
- Support up to 10 dual-width GPUs with maximum TDP of 600W. Each card supports PCIe 5.0 x16, providing extreme heterogeneous computing power.
- Support PCIe Switch x32 uplink, double that of the industry standard x16, meeting the high communication bandwidth requirements of CPU and GPU scenarios;
- Support up to 32 DDR5 memory modules with a maximum speed of 6400MT/s, increasing memory bandwidth by 33%.

### Flexible Architecture Rich Configuration

- Support both pass-through and switch architectures, and offers variety of GPU topologies to flexibly match the needs of different application scenarios.
- Extremely high expandability, supporting up to 15 standard PCIe slots + 1 OCP 3.0 network card slot.

### Extreme Heat Dissipation High-efficiency Power Supply

- Extremely optimized heat dissipation to meet the cooling requirements of 600W passive GPU.
- Power supply optimization, supporting PSU N+N redundancy

### Stable, Reliable and Intelligently Managed

- Key components of the system are all designed with redundancy and hot-swappability, and support tool-free disassembly and assembly, which improves the efficiency of fault maintenance and enhances the availability of the system.
- Integrates an intelligent management chip, provides an open management platform, and supports multiple management protocols such as IPMI2.0, Redfish, and SNMP;
- Support various management functions such as remote KVM, virtual media, critical component status monitoring, and abnormal alarms, realizing comprehensive remote system-level intelligent management.

# Technical Specifications

Function	Specifications		
Form	6U Rack		
CPU	2 * Intel® Xeon® 6 SP processors, TDP up to 350W		
GPU	Up to 10 full-height, full-length, double-width GPUs, with maximum TDP of 600W		
Memory	32 * DDR5 memory slots, RDIMM & 3DS RDIMM, with maximum speed of 6400MT/s 16 * MRDIMM memory modules, with maximum speed of 8000MT/s		
Storage Controller	Optional standard SAS HBA or RAID standard card, with optional M.2 RAID adapter card, supports Intel VMD and VROC.		
Local storage	<table border="1"> <tr> <td>12 * 3.5"hard drives (support 2.5") 4 * U.2 NVMe SSD Optional 2 M.2 slots, 2280 &amp; 22110, PCIe 5.0 x2/x4</td> <td>12 * 3.5"hard drives (support 2.5") 12 * U.2 NVMe SSD Optional 2 M.2 slots, 2280 &amp; 22110, PCIe 5.0 x4</td> </tr> </table>	12 * 3.5"hard drives (support 2.5") 4 * U.2 NVMe SSD Optional 2 M.2 slots, 2280 & 22110, PCIe 5.0 x2/x4	12 * 3.5"hard drives (support 2.5") 12 * U.2 NVMe SSD Optional 2 M.2 slots, 2280 & 22110, PCIe 5.0 x4
12 * 3.5"hard drives (support 2.5") 4 * U.2 NVMe SSD Optional 2 M.2 slots, 2280 & 22110, PCIe 5.0 x2/x4	12 * 3.5"hard drives (support 2.5") 12 * U.2 NVMe SSD Optional 2 M.2 slots, 2280 & 22110, PCIe 5.0 x4		
PCIe	<table border="1"> <tr> <td>11 PCIe 5.0 standard slots Optional 1 OCP 3.0 network card, PCIe 5.0 x4/x8 Optional 1 built-in PCIe 4.0 standard slot (for RAID cards)</td> <td>13 PCIe 5.0 standard slots Optional 1 OCP 3.0 network card, PCIe 5.0 x8/x16 Optional 2 built-in PCIe 4.0 standard slot (for RAID cards)</td> </tr> </table>	11 PCIe 5.0 standard slots Optional 1 OCP 3.0 network card, PCIe 5.0 x4/x8 Optional 1 built-in PCIe 4.0 standard slot (for RAID cards)	13 PCIe 5.0 standard slots Optional 1 OCP 3.0 network card, PCIe 5.0 x8/x16 Optional 2 built-in PCIe 4.0 standard slot (for RAID cards)
11 PCIe 5.0 standard slots Optional 1 OCP 3.0 network card, PCIe 5.0 x4/x8 Optional 1 built-in PCIe 4.0 standard slot (for RAID cards)	13 PCIe 5.0 standard slots Optional 1 OCP 3.0 network card, PCIe 5.0 x8/x16 Optional 2 built-in PCIe 4.0 standard slot (for RAID cards)		
I/O	Front: 2 USB 3.0 ports, 1 VGA port Rear: 1 serial port, 2 USB 3.0 ports, 1 VGA port, 1 RJ45 management port.		
Management	1 * 1Gbps RJ45 dedicated management port Integrated with the AST2600 BMC management chip, it supports IPMI 2.0, Redfish, SOL, KVM, virtual media, and other functions.		
Security	TPM/TCM security module, chassis opening intrusion detection, locks chassis top cover (tool-free). Dual flash redundancy design for BIOS/BMC; supports Intel SGX2.0 and TDX security technologies		
PSU	6 CRPS power modules, hot-swappable, and support 3+3 redundancy mode; Optional 2000W/2700W/3200W/3600W		
Fan	Optional rear fan module, supports N+1 redundancy, and supports hot-swapping.		
Chassis	W:448mm x H:263.5mm x D:869mm		
Operating Temperature	5°C - 35°C		
Operating Humidity	8% - 90%		
OS	Mainstream operating systems such as Ubuntu, Red Hat Enterprise, and Windows Server.		