

GPTrack.ai GH200 624GB Manual

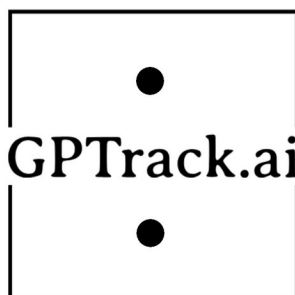
Version 1.4 / December 2024

2024 GPTrack.ai UG (limited)

For the latest information and updates please see: <https://gptrack.ai>

Table of contents

| | |
|----------------------------|----|
| Quickstart..... | 2 |
| System specifications..... | 2 |
| Package Contents | 3 |
| System Rear I/O..... | 4 |
| Power Sub-System..... | 5 |
| BMC..... | 6 |
| Firmware..... | 6 |
| Operating system..... | 6 |
| Drivers..... | 7 |
| Software..... | 8 |
| Nvidia Resources..... | 9 |
| User Guide..... | 10 |



Quickstart:

Username and password for BMC:

Username: root

Password: OpenBmc (zero!!!)

Username and password for preinstalled OS (optional):

Username: x

Password: xz

1.) Install OS

Ubuntu Server for ARM: <https://cdimage.ubuntu.com/releases/24.04/release/ubuntu-24.04-live-server-arm64+largemem.iso>

2.) Install Drivers

```
curl -O https://us.download.nvidia.com/tesla/565.57.01/NVIDIA-Linux-aarch64-565.57.01.run
```

```
sudo apt install build-essential
```

```
sudo sh NVIDIA-Linux-aarch64-565.57.01.run
```

System specifications:

Nvidia GH200 Grace Hopper Superchip

72-core Nvidia Grace CPU

Nvidia H200 Hopper Tensor Core GPU

480GB of LPDDR5X memory with EEC

144GB of HBM3e memory

624GB of total fast-access memory

NVlink-C2C: 900 GB/s of bandwidth

Programmable from 400W to 1000W TDP (CPU + GPU + memory)

2x High-efficiency 2000W PSU

2x PCIe gen4 M.2 slots on board

2x PCIe gen5 2.5" drive slots (NVMe) without Bluefield-3

4x PCIe gen5 2.5" drive slots (NVMe) with Bluefield-3

3x FHFL PCIe Gen5 x16

1x USB 3.2 port (mini USB hub included: 1x USB 3.0, 2x USB 2.0)

2x RJ45 10GbE ports

1x RJ45 IPMI port

1x Mini display port

1x Micro USB port

Halogen-free LSZH power cables

Stainless steel cage nuts

Air-cooled 6x60mm fans
Rail kit
2U 440 x 88 x 900 mm (17.3 x 3.5 x 35.4")
34 kg (75 lbs)

TPM (optional, SPI mode)
ACPI compliance, S0, S5 support

System rating:
100-120Vac, 50/60Hz, 10A
200-240Vac, 50/60Hz, 10A

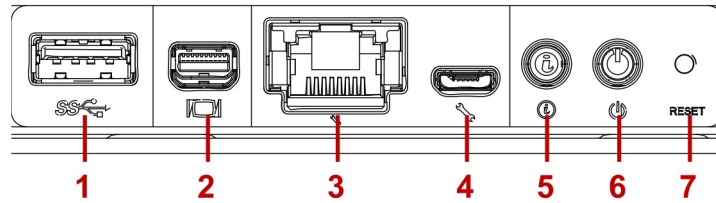
System management: IPMI v2.0 Compliant, on board "KVM over IP" support, Dedicated GbE management NIC port from PHY RTL8211FS to BMC

Operating environment:
Operating temperature: 0°C to 35°C (41°F to 95°F) at 1829M and up to 50°C at 0M. (No performance drop)
Non-operating temperature: -40°C to 70°C (-40°F to 158°F)
Operating relative humidity: 5% to 85%RH
Non-operating relative humidity: 50% to 93%RH

Package Contents







1x GH200 624GB system
2x Power cord
4x Cage nut
1x Rail kit

System Rear I/O

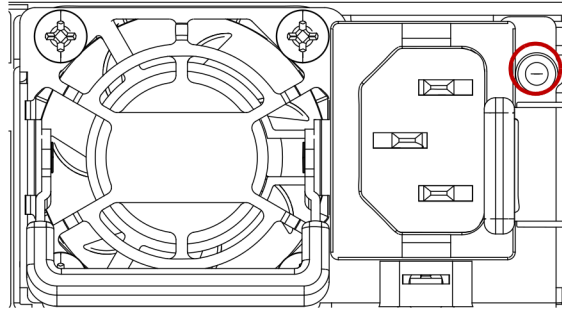


System Rear I/O

System Rear I/O Definition

| No. | ICON | NAME | DESCRIPTION |
|-----|---|---------------------|--|
| 1 |  | USB 3.0 port | Connect to USB device Note: The USB device connected must be: No bigger than 17.7 mm (W) x 9.5 mm (H) to avoid interference with other ports. |
| 2 |  | Mini-DP connector | Maximum display resolution: 1920x1200 32bpp@60Hz (reduced blanking) |
| 3 |  | Dedicated NIC | Dedicated RJ45 connector |
| 4 |  | Micro USB port | Transmit in serial signal for debug or terminal concentrator |
| 5 |  | Location button/LED | Press to trigger on/off ID LED. Blue blinking - Identifier; Off - Normal. |
| 6 |  | Power button/LED | Press to power on the system. Press and hold for four seconds to power off the system. Blue blinking - System power off; On - System power on. |
| 7 | | Reset button | Press to reset the system |

Power Sub-System



PSU to Mainboard Module Description

Two power supply unit are supplied in the system. A secondary PSU is available for redundancy functionality.

Redundant AC Power Supply Units by Model

| PSU | AC INPUT |
|-------------------------------|--------------------------------|
| (2) 2000W 73.5mm Platinum PSU | 100-240VAC 50/60Hz, AC support |

Power Supply Unit LED

| PSU LED COLOR | DESCRIPTION |
|---------------------|---|
| Green On | Output ON and OK |
| Green Blinking | AC present / Only VSB ON |
| Amber On | AC Lost; with a second PSU in parallel still with AC input power. PSU critical event causing a shutdown; failure, OCP, OVP, Fan Fail. |
| Amber Blinking @1Hz | PSU warning events where the PSU continues to operate; high temp, high power, high current, slow fan, UV. |

BMC

How to connect to BMC:

Connect network cable to the dedicated IPMI port. Get the IP.

Open browser and enter IP: <https://192.168.178.x>

Standard username and password for BMC

username: root

password: OpenBmc (zero!!!)

Firmware Update

Download the firmware from <https://gptshop.ai>

Select update firmware in BMC to update (use tar file, do not extract).

Operating system

Standard username and password for preinstalled OS (optional):

username: x

password: xz

Install OS yourself:

Any ARM linux distribution with kernel > 6.5 should work just fine.

Example Ubuntu (standard):

Download the version you would like to have (64k kernel is highly recommended).

Ubuntu Server for ARM: <https://ubuntu.com/download/server/arm>

Intall with USB stick or over network via BMC (slower).

With older versions (not recommended) you might need to select the HWE kernel. Using the newest 64k kernel is highly recommended.

There are special Nvidia kernels for Ubuntu 22.04: [linux-nvidia-64k-hwe](#)

If you want to install a Nvidia kernel do it like this:

```
dpkg --list|grep linux-image
```

```
sudo apt purge linux-image-xxx (xxx = version number)
```

```
dpkg --list|grep linux-headers
```

```
sudo apt purge linux-headers-xxx (xxx = version number)
```

```
dpkg --list|grep linux-modules
```

```
sudo apt-get purge linux-modules-xxx (xxx = version number)
```

```
sudo apt update
```

```
sudo apt install linux-nvidia-64k-hwe-22.04-edge
```

```
sudo reboot
```

Drivers

There are two main ways to install the Nvidia drivers

1.) Via Nvidia driver installer

Download Nvidia GH200 drivers: <https://www.nvidia.com/Download/index.aspx?lang=en-us>

Select product type "data center", product series "HGX-Series" and operating system "Linux aarch64"

Copy the download url (may change with version and over time)

example:

```
curl -O https://us.download.nvidia.com/tesla/xxx/NVIDIA-Linux-aarch64-xxx.run (xxx = version number)
```

```
sudo apt install build-essential
```

```
sudo sh NVIDIA-Linux-aarch64-xxx.run -m=kernel-open (xxx = version number)
```

2.) Via package installer (example Ubuntu)

Check version if already installed:

```
cat /proc/driver/nvidia/version
```

List possible driver for install

For desktop:

```
sudo ubuntu-drivers list
```

or, for servers:

```
sudo ubuntu-drivers list --gpgpu
```

Automatic install (desktop):

```
sudo ubuntu-drivers install
```

Installing the drivers on servers

```
sudo ubuntu-drivers install --gpgpu
```

You will also want to install the following additional components:

```
sudo apt install nvidia-utils-xxx-server (xxx = version number)
```

Alternatively, manual driver installation (using APT)

```
sudo apt install nvidia-driver-xxx (xxx = version number)
```

Software

Nvidia CUDA

There are two ways to install the Nvidia CUDA

1.) Nvidia installation guide:

<https://docs.nvidia.com/cuda/cuda-installation-guide-linux/contents.html>

Downlad from:

https://developer.nvidia.com/cuda-downloads?target_os=Linux&target_arch=arm64-sbsa

2.) Via package intaller (example Ubuntu)

```
sudo apt install nvidia-cuda-toolkit
```

```
sudo apt install nvidia-cuda-dev (optional)
```

Nvidia resources

Official Nvidia GH200 Manual: <https://docs.nvidia.com/grace/#grace-hopper>

Official Nvidia Grace Manual: <https://docs.nvidia.com/grace/#grace-cpu>

Official Nvidia Grace getting started: <https://docs.nvidia.com/grace/#getting-started-with-nvidia-grace>

