Firefly

10-node 1000T Cluster Server

CSB1-N10NOrinNano

CSB1-N10NOrinNX

Product features

1000TOPS powerful computing power

The total computing power of the server is up to 1000TOPS, with 10 built-in distributed computing nodes + 1 control node, which can run mainstream modern AI models: ROS robot models, etc.

Powerful multi-channel video processing capabilities

It supports up to 10 channels of H.265 8K@30fps, 20 channels of H.265 4K@60fps video decoding, and 10 channels of H.265 4K@60fps video encoding.

The private deployment of large models

Large language models: Support Ollama local large model deployment framework and the private deployment of ultra-large-scale parametric models: LLaMa3 and Phi-3 Mini.
Vision models: Support EfficientVIT, NanoOWL, NanoSAM, SAM and TAM.
AI painting: Support ComfyUI graphical deployment framework and the private deployment of Flux, Stable Diffusion image generation model in the AIGC field.

Multiple deep learning frameworks

Supports multiple deep learning frameworks accelerated by cuDNN, including PaddlePaddle, PyTorch, TensorFlow, MATLAB, MxNet, Caffe2, Chainer and Keras, as well as custom operator development. Docker containerization technology is supported.

Product features

Secure and high-speed network communication

It has 2×10 Gigabit Ethernet ports, 2×Gigabit Ethernet ports, and 1×BMC management network interface, so that network communication has a higher speed.

Standard 1U rack server

Highly dense and tightly deployed, with 10 computing modules per unit; The standard 1U rack server chassis design perfectly matches most types of cabinets in the data center.

Equipped with BMC management system

Equipped with BMC intelligent management system, it can easily complete real-time monitoring, software configuration, hardware management, troubleshooting, system upgrade, and can provide secondary development.

Wide range of application scenarios

It is widely used in edge computing, robotics, large model localization, smart city, smart healthcare, smart industry and other industries.
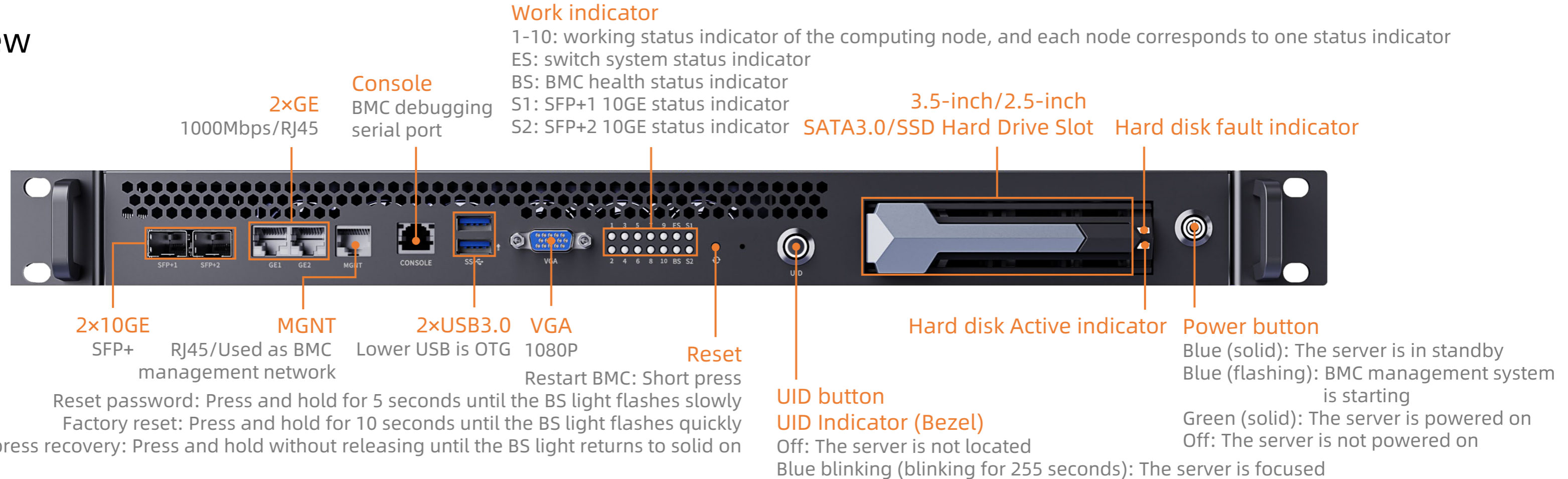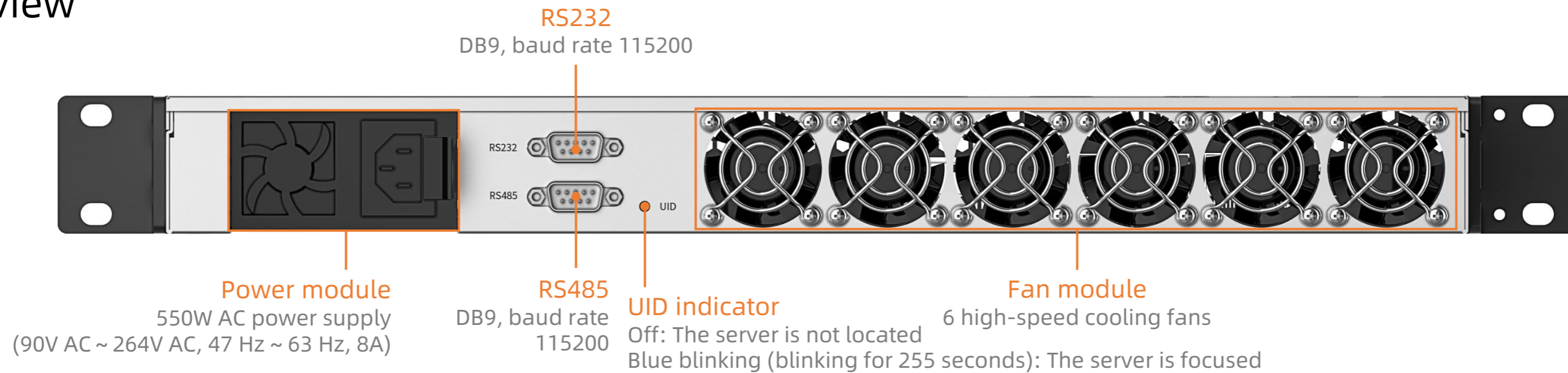
Specifications

		CSB1-N10NOrinNano	CSB1-N10NOrinNX
Technical Specifications	Server form	1U rack-mounted computing power server	
	Framework	ARM architecture	
	Number of nodes	10 distributed computing nodes (up to 80 ARM cores) + 1 control node	
	Compute nodes	Hexa-core 64-bit processor NVIDIA Jetson OrinNano, main frequency up to 1.5GHz	Octa-core 64-bit processor NVIDIA Jetson OrinNX, main frequency up to 2.0GHz
	Control nodes	Octa-core 64-bit processor RK3588, main frequency up to 2.4GHz, the highest computing power is 6TOPS	
	AI computing power	400TOPS (INT8)	1000TOPS (INT8)
	RAM	8GB LPDDR5 × 10 (Number of compute nodes)	16GB LPDDR5 × 10 (Number of compute nodes)
	Storage	256GB (2242 PCIe NVMe SSD, the server is internally assembled)	
	Storage Expansion	3.5-inch/2.5-inch SATA3.0/SSD hard drive slot × 1 (BMC can directly operate the hard drive, and computing child nodes can indirectly access the hard drive through the network sharing method provided by BMC)	
	Power	550W AC power supply (Input: 90V AC~264V AC, 47 Hz~63 Hz, 8A) (Hot swappable not supported)	
	Fan module	6 high-speed cooling fans	
Physical Specifications	Size	420.0mm(L) × 421.3mm(W) × 44.4mm(H)	
	Installation requirements	IEC 297 Universal Cabinet Installation: 19 inches wide and 800 mm deep and above Retractable slideway installation: The distance between the front and rear holes of the cabinet is 543.5mm~848.5mm	
	Full weight	Server net weight: 8.1kg, total weight with packaging: 10.3kg	
	Environment	Operating Temperature: 0ºC ~ 45ºC, Storage Temperature: -40ºC ~ 60ºC, Operating Humidity: 5% ~ 90%RH(non-condensing)	
Software Specifications	BMC	The BMC management system is integrated with the web-based management interface, supporting Redfish, VNC, NTP, monitoring advanced and virtual media, and the BMC management system can be redeveloped.	
	Large model	Robot model: ROS robot model is supported. Large language models: Support Ollama local large model deployment framework, which can be used for natural language processing, code generation, and assistance scenarios. Support the private deployment of ultra-large-scale parametric models under the Transformer architecture, such as LLaMa3 and Phi-3 Mini. Large visual models: Support the privatization deployment of large visual models such as EfficientVIT, NanoOWL, NanoSAM, SAM and TAM. AI Painting: Support ComfyUI graphical deployment framework, which can be used for scenarios such as image restoration, image style conversion, and image synthesis. Supports the private deployment of Flux, Stable Diffusion and Stable Diffusion XL image generation model in the AIGC field.	
	Deep learning	Supports multiple deep learning frameworks accelerated by cuDNN, including PaddlePaddle, PyTorch, TensorFlow, MATLAB, MxNet, Caffe2, Chainer and Keras. Supports custom operator developmen. Docker containerization: Docker containerization technology is supported, which can be easily used for image deployment.	
	AI software stack	The NVIDIA Jetson Orin series delivers powerful AI compute power, massive unified memory, and a comprehensive software stack to power the latest generative AI applications. It enables fast inference on any generative AI model powered by the Transformer architecture, enabling superior edge performance on MLPerf.	
Interface Specifications	Internet	2 × 10G Ethernet (SFP+), 2 × Gigabit Ethernet (RJ45), 1 × Gigabit Ethernet (RJ45, MGNT is used as BMC management network)	
	Console	1 × Console (RJ45, BMC debug serial port, baud rate 115200)	
	Display	1 × VGA (maximum resolution 1080P, BMC management display)	
	USB	2 × USB3.0 (The lower USB is USB3.0 OTG, and the BMC can be upgraded OTG by using a USB flash drive)	
	Button	1 × Reset, 1 × UID, 1 × Power button	
	Other interfaces	1 × RS232 (DB9, baud rate 115200),1 × RS485 (DB9, baud rate 115200)	

Interface description

Front view

2×GE
1000Mbps/RJ45

Console
BMC debugging
serial port

Work indicator
1-10: working status indicator of the computing node, and each node corresponds to one status indicator
ES: switch system status indicator
BS: BMC health status indicator
S1: SFP+1 10GE status indicator
S2: SFP+2 10GE status indicator

3.5-inch/2.5-inch
SATA3.0/SSD Hard Drive Slot

Hard disk fault indicator

2×10GE
SFP+

MGNT
RJ45/Used as BMC
management network

2×USB3.0
Lower USB is OTG

VGA
1080P

Reset
Restart BMC: Short press

Reset password: Press and hold for 5 seconds until the BS light flashes slowly
Factory reset: Press and hold for 10 seconds until the BS light flashes quickly
Accidental press recovery: Press and hold without releasing until the BS light returns to solid on

UID button
UID Indicator (Bezel)
Off: The server is not located
Blue blinking (blinking for 255 seconds): The server is focused

Hard disk Active indicator

Power button
Blue (solid): The server is in standby
Blue (flashing): BMC management system
is starting
Green (solid): The server is powered on
Off: The server is not powered on

Rear view

RS232
DB9, baud rate 115200

Power module
550W AC power supply
(90V AC ~ 264V AC, 47 Hz ~ 63 Hz, 8A)

RS485
DB9, baud rate
115200

UID indicator
Off: The server is not located
Blue blinking (blinking for 255 seconds): The server is focused

Fan module
6 high-speed cooling fans

Dimension

421.3mm

44.4mm

420.0mm

T-CHIP INTELLIGENCE TECHNOLOGY

Contact Us
(+86)18688117175

E-mail
global@t-firefly.com

Website
https://en.t-firefly.com/

Address
Room 2101, Hongyu Building, #57 Zhongshan 4Rd, East District, Zhongshan, Guangdong, China.