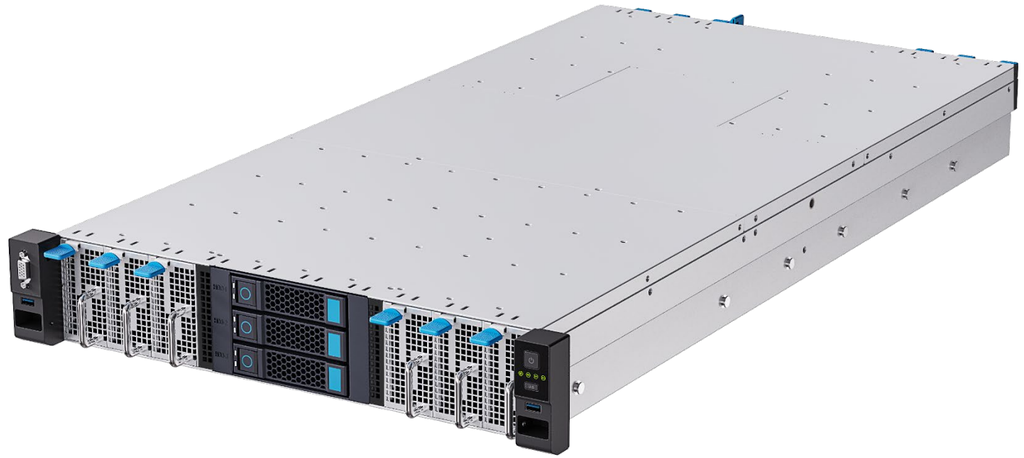
CSD2-N96 Power Server

- RK3588S, RK3576
- BM1688
- QCS8550
- SpacemiT K3



V1.0 2026-5-22

FIREFLY TECHNOLOGY

Product features

Supports mainstream SoCs including Qualcomm and RockChip

It features a wide range of server computing processors, with optional mainstream hardware platforms such as Qualcomm, RockChip, SOPHGO and SpacemiT, suitable for diverse computing-intensive application scenarios.

Fully configured with 96 computing nodes

When fully configured, the device can deploy 96 computing nodes. Each node supports parallel deployment of 5–10 containers based on business requirements, and the whole unit can virtually host 480–960 system containers.

Limited impact from single-point failures

Each computing blade module is equipped with 8 core boards. Maintenance on a single faulty core board only affects the corresponding 8 service nodes, which is superior to the conventional industry standard where 16–20 service nodes are impacted.

Full-domain hot-swappable modular design

The whole machine adopts a full-domain hot-swappable modular architecture, integrating BMC network module, switching network module, power supply module and 12 computing blade modules. All modules support online hot-swap.

Product features

Supports multi-instance Android system operation

Adopting an Android container-level isolated multi-instance architecture, it efficiently reuses SoC hardware resources through lightweight virtualization scheduling to enhance computing power and overall system utilization.

Peak aggregated bandwidth up to 80Gbps

Equipped with 8×10G SFP+ ports, it delivers an aggregated peak bandwidth of up to 80Gbps to meet high-bandwidth scenario requirements. An independent BMC management network interface separates the management network from the service network, ensuring secure and reliable network communication.

Supports three 3.5-inch SATA3.0 HDD/SSD

Boasting three 3.5-inch (or 2.5-inch) drive bays, the device supports SATA3.0 HDD/SSD expansion for effortless TB-level large capacity storage. With hot-swappable functionality for quick drive replacement, it perfectly meets one-stop deployment needs for file management, data backup, and video surveillance scenarios.

Wide application scenarios

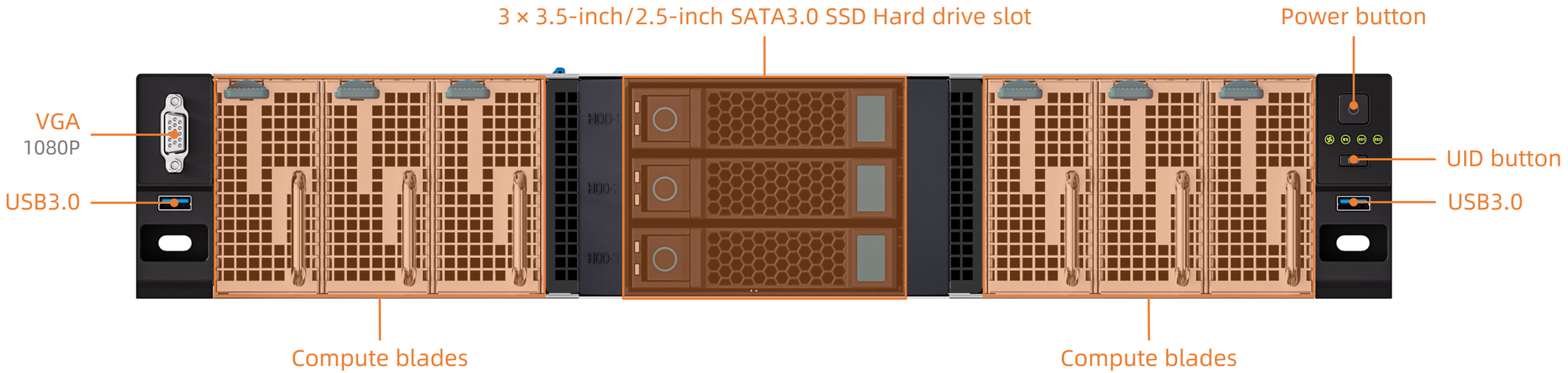
Widely used in edge computing, on-premises large model deployment, smart city, smart healthcare, smart industry, intelligent security and other fields.

Specifications

		QCS8550	RK3588S	RK3576	BM1688	K3
Technical Specifications	Product model	CSD2-N96Q8550	CSD2-N96R3588S	CSD2-N96R3576	CSD2-N96S1688	CSD2-N96SPK3
	Server form	2U rack-mounted computing power server				
	Architecture	ARM architecture				RISC-V architecture
	Number of nodes	12 compute blades (96 distributed compute nodes) + 1 control node				
	Compute nodes	Octa-core 64-bit processor Qualcomm QCS8550, up to 3.36GHz	Octa-core 64-bit processor RK3588S, up to 2.4GHz	Octa-core 64-bit processor RK3576, up to 2.2GHz	Octa-core 64-bit processor BM1688, up to 1.6GHz	Octa-core 64-bit processor SpacemiT Key Stone K3, up to 2.4GHz
	Video encoding	8K@30fps/4K@120fps H.265/H.264	H.265&H.264: 1×8K@30fps, 16×1080P@30fps	H.265&H.264: 1×4K@60fps	H.265&H.264: 10×1080P@30fps	4K@90fps H.265/H.264
	Video decoding	8K@60fps/4K@240fps H.265/H.264/VP9/AV1	8K@60fps/4K@120fps (H.265/VP9/AVS2) 8K@30fps (H.264/AVC/MVC) 30×1080P@30fps (H.265& H.264)	1×4K@120fps (H.265/HEVC,VP9,AVS2,AV1) 1×4K@60fps (H.264/AVC)	H.265&H.264: 10×1080P@30fps	4K@180fps H.265/H.264/VP9
	Control nodes	Octa-core 64-bit processor RK3588, main frequency up to 2.4GHz, the highest computing power is 6TOPS				
	AI computing power	4608TOPS (48T × 96, INT8)	576TOPS (6T × 96, INT8)	576TOPS (6T × 96, INT8)	1536TOPS (16T × 96, INT8)	5760TOPS (60T× 96, SquareINT4)
	RAM	16GB LPDDR5X × 96	16GB LPDDR5 × 96 (4/8/16/32GB)	8GB LPDDR4/LPDDR5 × 96 (4/8/16GB)	8GB LPDDR4 × 96 (4/8/16GB)	32GB LPDDR5 × 96 (8/16/32GB)
	Storage	256GB UFS4.0 × 96	256GB eMMC × 96 (16/32/64/128/256GB)	64GB eMMC × 96 (16/32/64/128/256GB)	32GB eMMC × 96 (16/32/64/128/256GB)	128GB UFS2.2 × 96
	Storage expansion	3.5-inch/2.5-inch SATA3.0 SSD hard drive slot × 3 (Supports hot swapping; 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	2 × 1300W hot-swappable power supplies, 1+1 redundancy support				
	Fan module	14 high-speed cooling fans				
Physical Specifications	Size	Standard 2U rack servers: 495.60mm × 928.51mm × 88.80mm				
	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				
	Environment	Operating Temperature: 0°C ~ 30°C, Storage Temperature: -40°C ~ 60°C, Operating Humidity: 5% ~ 80%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 language models	All models support private deployment of ultra-large-scale parameter models under the Transformer architecture, such as large language models including Deepseek-R1 Series, Gemma Series, Llama Series, ChatGLM Series, Qwen Series, Phi Series, etc.				
	Visual large model	K3: Supports private deployment of all vision large models QCS8550: Supports private deployment of vision large models including Qwen2.5-VL, InternVL3, etc.				
	AI Painting	K3: Supports private deployment of all image generation models QCS8550: Supports private deployment of the Stable Diffusion image generation model				
	Deep learning	All models: Support traditional network architectures such as CNN, RNN, LSTM, and support various deep learning frameworks such as TensorFlow, PyTorch, PaddlePaddle, ONNX, and Caffe. Support custom operator development and Docker containerization management technology				
Interface Specifications	Internet	8 × 10Gbps SFP+, 1 × Gigabit Ethernet (RJ45, MGMT 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	3 × USB3.0, 1 × Type-C (OTG)				
	Button	1 × Power, 1 × UID, 1 × Recovery, 1 × Reset				

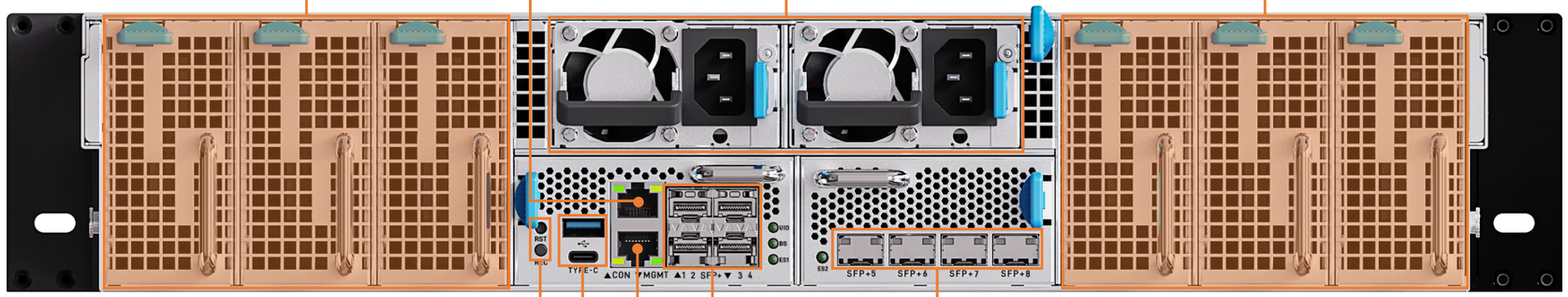
Interface description

Front view

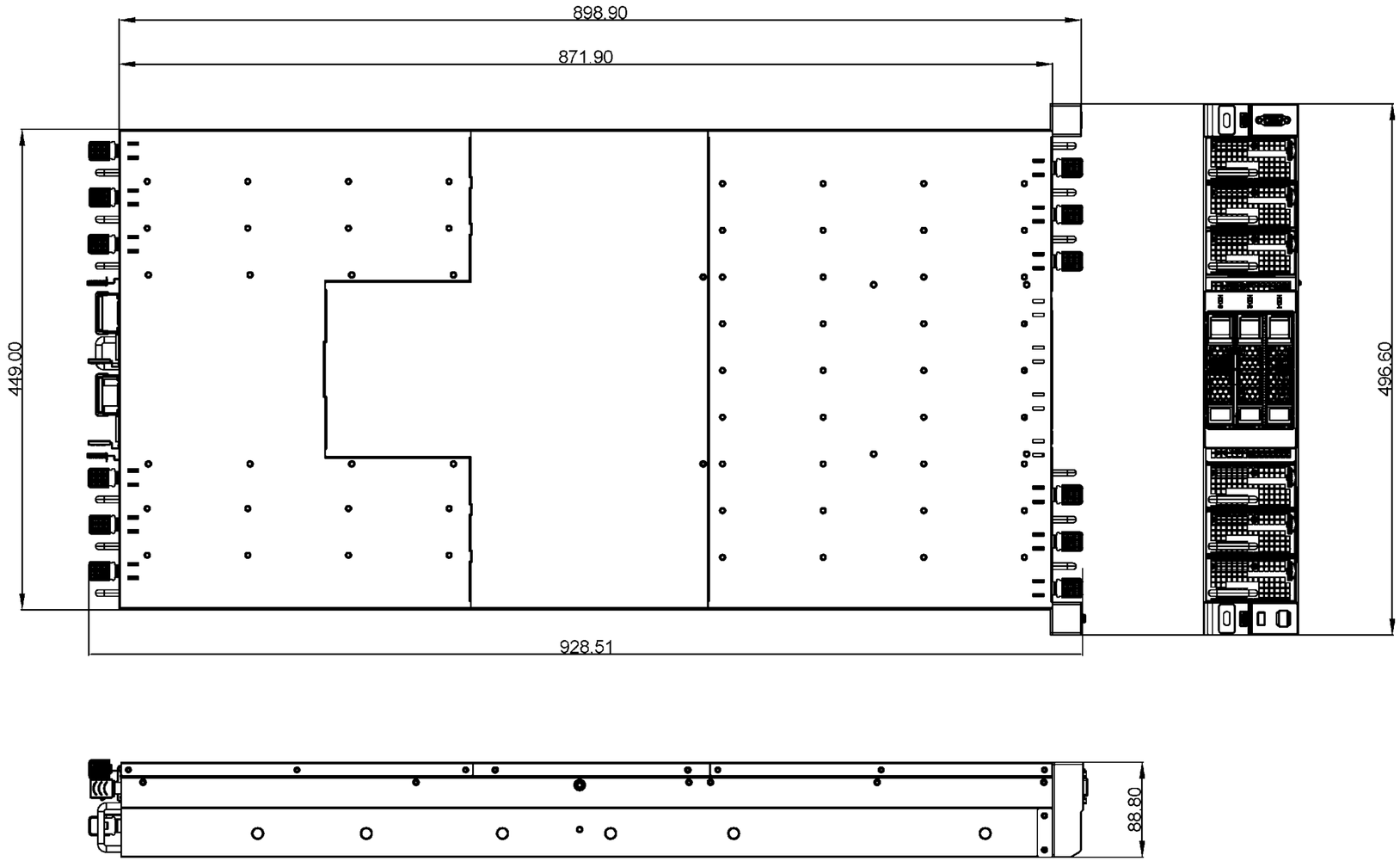


Interface description

Rear view



Dimension



FIREFLY 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.