

## The Heart of Al applications: GPU versus CPU.

#### CPU



#### **Functions**

CPU's primary purpose is to perform instructions given by a software program, such as arithmetic, logic, and I/O controls.



#### Components

The CPU encompasses an MMU, Cache memory is faster than RAM and is placed on the processor's chip, causing it to be closer to the CPU.



#### **Performance**

CPU performance can be measured by several characteristics, including:

Clock speed: The number of cycles of execution of a program that can fit into a second

A number of cores: CPUs are classified based on the number of cores they have: single-core, dual-core, quad-core, Hexa-core, eight-core, and ten-core machines.

#### **GPU**



#### **Functions**

GPUs are designed with hundreds of individual cores, the ability of which to support thousands of threads at one time and thus optimize the product.



#### **Video Memory**

Aimed at controlling big bandwidth size from graphical-use applications, it has been specially planned for this purpose.



#### **Cooling Systems**

Some models of GPU are equipped with fans or exhaust systems to avoid heating and speed up the process to the level of efficient cooling.



#### **Low Latency**

Some specific GPUs available today are configured to measure/improve system latency to focus on performance enhancement, improved target-acquisition and reaction times.

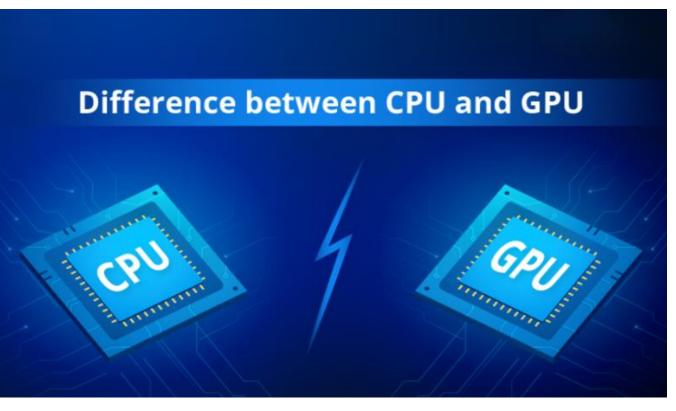


#### **Performance Drivers**

GPUs are the primary performance drivers in HEAVY.Al and CPUs play a supportive role.



## The Heart of Al applications: GPU versus CPU.

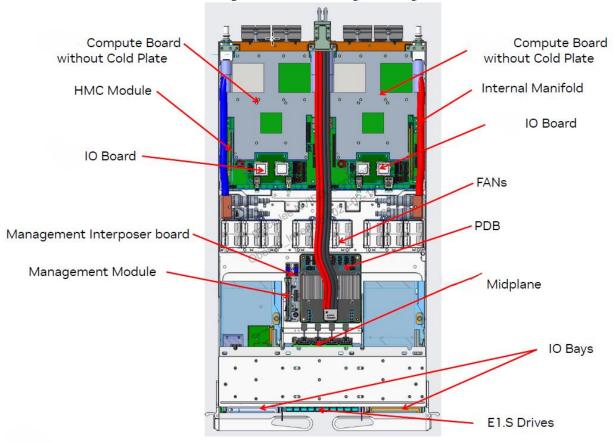


- The Central Processing Unit (CPU) and Graphics Processing Unit (GPU) both have unique computing strengths. CPU and GPU working together can optimize performance in many systems. The CPU handles general-purpose tasks and system management, on the other hand, the GPU takes on specific, computation-heavy tasks. This CPU/GPU combination offers better efficiency and faster processing in various applications.
- CPUs are crucial for tasks related to sequential processing and complex algorithmic calculations that need to be done one by one.
- GPUs are designed for parallel processing so they are ideal for training AI models. AI training involves performing similar operations on multiple data samples simultaneously. GPUs provide the necessary computational power for tasks like neural networks, accelerated AI operations, and traditional AI inference and training algorithms.



## **GB-200** drives Al applications: GPU + CPU.

### **Compute Tray Layout**







## **Delta OCP Solution**

CLOUD DATA CENTER

0.0.0.0.0



The future of white space infrastructure solutions for AI applications:

**ORV3 Rack** 

33 KW PS System

18 KW PS System

**AALC** 

15kW/27.5kW Bat. Backup System

Liquid-to-Air or Liquid-to-Liquid
CDU







**66kW Power Supply System**U, 55 kW 50 VDC with 5.5 kW PSU modules



**33kW Power Supply System**U, 27.5 kW 50 VDC with 5.5 kW PSU modules



18kW Power Supply System
U. 15 kW 50 VDC with 3 kW PSU modules







### **OCP Solution**

(ORV3+AALC+CDU)

**TRENDS** 

**REQUIREMENTS** 

MARKET POTENTIAL

Liquid Cooling with the Open Compute Project (OCP) is a MUST-HAVE for Al applications

Building region competence over the OCP related products in product training, product information, capability to provide service and technical expertise.

The share of OCP + Liquid cooling is expected to grow rapidly with CAGR over 23% in the next 5 years.

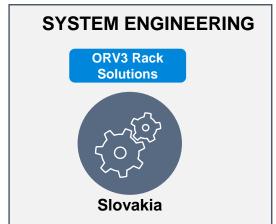
Moving forward colocation providers are expected to adopt this new technology overtime.

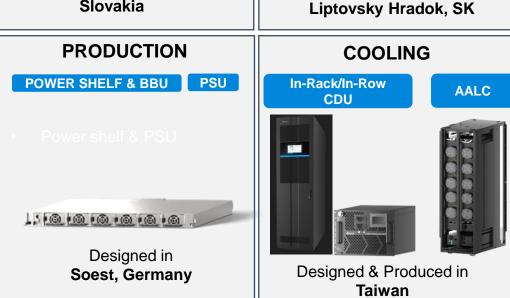


## **OCP RACK** | New Establishment

### Plans to Meet HPC & Al Demands









**ASSEMBLY** 

**RACK** 

Racks Assembled in

## **GPU Load Characteristics**

**GPU Electrical Data Peak Processing** 

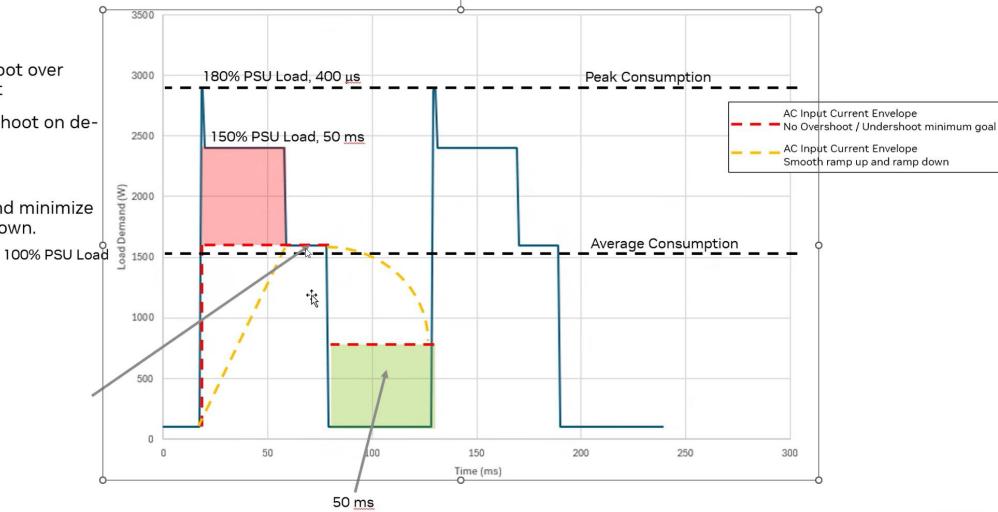
#### **Red Line: Minimum**

 Need to have 0% overshoot over average AC input current

Need to not have undershoot on deloading.

#### Orange Line: Target

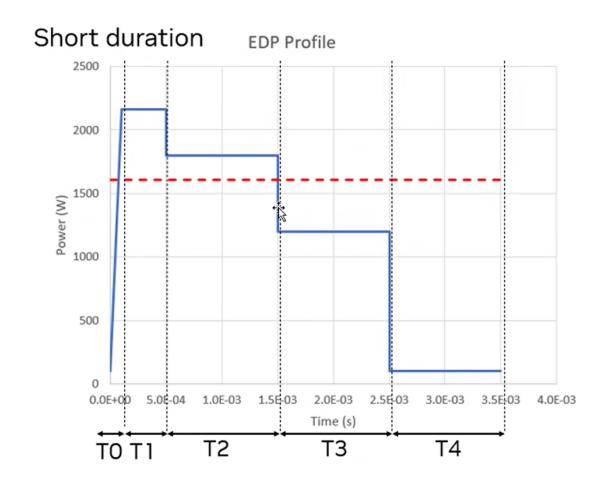
1. Maximize holdup time and minimize ramp rate both up and down.





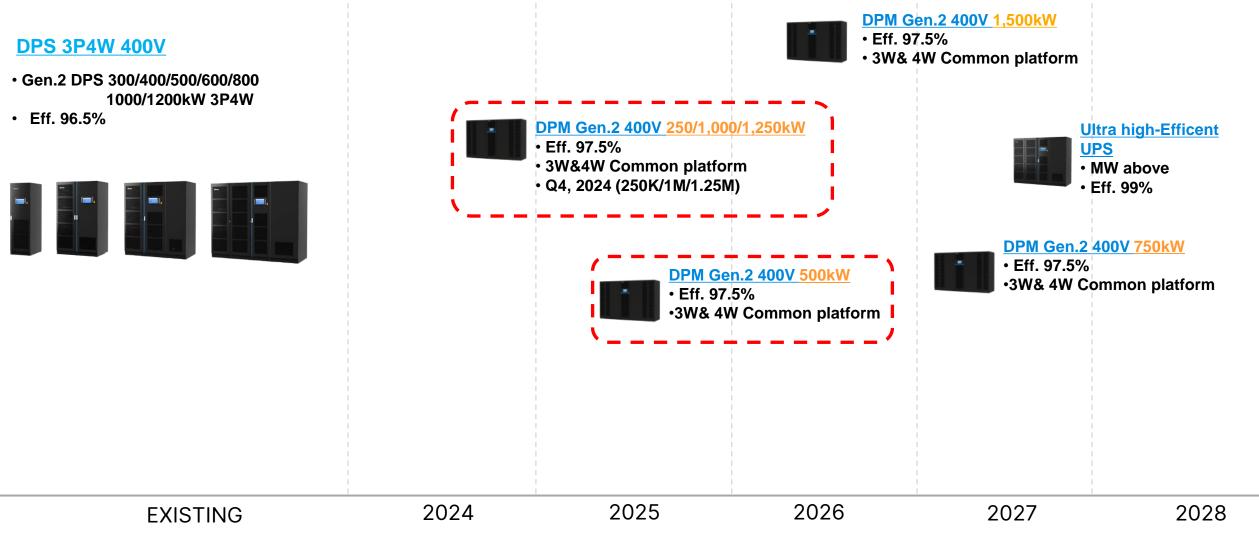
### **GPU Load Characteristics**

- The GPU load profile can be broken down by 5 different time intervals
- T0 Initial turn on slew
- T1 settling time before reaching EDPp1 –400us maximum
- T2 EDPp1 peak limiting target value 50ms maximum
- T3 EDPc unlimited duration
- T4 waiting for next cycle unlimited duration





## **Large UPS for Al Product Roadmap**





# **DPM G2 General Specifications**

		DPM Gen 2
Power Rating	Ready	250/ 1000/ 1250 kW
	Upcoming	500/ 750/ 1500/ 1750 kW
Input	Nominal Voltage	380/400/415 Vac, 3P3W+PE/ 3P4W+PE
	THDi	<3%
	Power Factor	<0.99
Output	Voltage	380/400/415 Vac, 3P3W+G/ 3P4W+PE
	Voltage Regulation	±1% (static) VFI-SS-111 (dynamic)
	THDv	< 1% (linear load)
	Overload Capability	≦110%: continues, ≦125% : 10 minutes ≦150% : 1 minute, >150% : 1 sec
Battery	Battery Type	Lithium-ion, VRLA, Ni-Zinc
	Nominal Voltage	480V
	Quantity	34-35 <sup>1</sup> , 36-46pcs (VRLA 12V)
Efficiency	Online Mode	Up to 97.3%
Parallel Configuration		Up to 8 units
Conformance	Safety	CE, UKCA
	EMC	IEC 62040-2
	Performance	IEC 62040-3







## **UPS Li-ion Battery Product Roadmap**

#### **UBH3 35.5 kWh**



- Cell: NMC Delta P140
- Full front access
- UL9540A/IEC62619
- Depth 625mm/against the wall
- 3W battery systems
- Oct. 2023

#### UZR3 31.0 kWh/62.1 kWh



- Cell: NMC Delta P140
- 2string 60Ah inside
- UL9540A/IEC62619
- 2W battery systems
- Oct, 2023

### NMC Technology

### LFP technology

#### LZM/LBM 43.5kWh



- Cell: LFP 27Ah
- UL1973/UL9540A/ IEC62619/CE
- 2W/3W battery systems
- Q4, 2024
- Max. Power 200kW

#### **LZH** 47.1kWh (<10min)



UL1973/UL9540A/

IEC62619/CE

- 2W battery systems
- Q1, 2025
- Max. Power 400kW

#### **LZT** 180kWh (>2h)

- Cell: LFP 314Ah
- UL1973/UL9540A/ IEC62619/CE
- 2W/batt. System
- Conceptual
- Max. Power 50kW

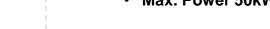
#### **LZE 51.2kWh (1h)**



- Cell: LFP 100Ah
- UL1973/UL9540A/

#### IEC62619/CE

- 2W/batt. System
- Q1. 2025
- Max. Power 50kW



2024 2026 **EXISTING** 2025



