AI ACCELERATOR

PNY | **PHISON**

aiDAPTIV+
Scaling Boundaries, Not Budgets

EASE OF USE
aiDAPTIV+ allows you to spend more time training your data, not your team of engineers.

COST AND ACCESSIBILITY
aiDAPTIV+ leverages cost-effective NAND flash to increase access to large-language model (LLM) training with commodity workstation hardware.

PRIVACY
aiDAPTIV+ workstations allows you to retain control of your data and keep it on premises.

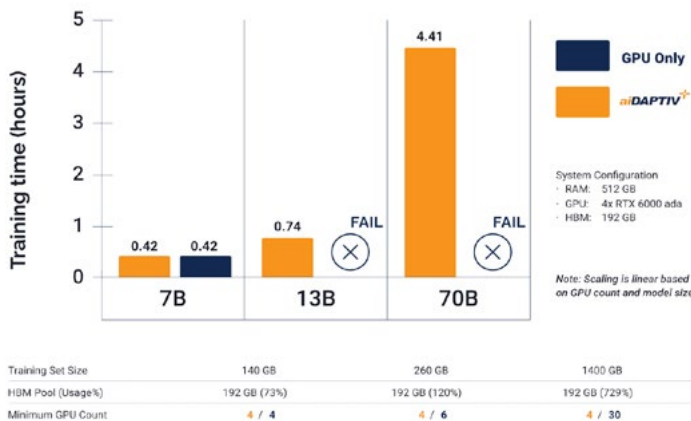## HYBRID SOLUTION BOOSTS LLM TRAINING EFFICIENCY

aiDAPTIV+ is a hybrid software / hardware solution for today's biggest challenges in LLM training. A single local workstation or server from one of our partners provides a cost-effective approach to LLM training, up to Llama-3 70B and Falcon 180B.

DROP-IN SOLUTION FOR PYTORCH

Experience seamless integration with the benefits offered by this system. It features a transparent drop-in function that eliminates the need to modify your AI application. You can effortlessly reuse existing hardware or add nodes as needed. System integrators have access to AI100E drives, middleware library licenses, and full support to facilitate smooth system integration.

SEAMLESS INTEGRATION WITH GPU MEMORY

The optimized middleware extends GPU memory capacity by utilizing up to 2x 2TB aiDAPTIVCache to support a 70B model with low latency. Additionally, the high endurance feature offers an industry-leading 100 DWPD over 3 years, utilizing pSLC NAND with an advanced NAND correction algorithm.

	7B	13B	70B
Training Set Size	140 GB	260 GB	1400 GB
HBM Pool (Usage%)	192 GB (73%)	192 GB (120%)	192 GB (729%)
Minimum GPU Count	4 / 4	4 / 6	4 / 30

Note: Scaling is linear based on GPU count and model size

System Configuration
- RAM: 512 GB
- GPU: 4x RTX 6000 ada
- HBM: 192 GB

UNLOCK LARGE MODEL TRAINING

Until aiDAPTIV+, small and medium-sized businesses have been limited to small, imprecise training models with the ability to scale beyond Llama-2 7b. aiDAPTIV+ solution enables significantly larger training models, giving you the opportunity to run workloads previously reserved for data-centers.

PRODUCT INFORMATION

PNY Part Number (Drives sold in pairs*)	1TB: **M280AI3140-1TB-PB**
	2TB: **M280AI3140-2TB-PB**
Service & Support	3-Year Limited Warranty
Usable Capacities (IDEMA)	1TB, 2TB
UPC	1TB: 751492792002
	2TB: 751492792019
Box Dimensions	5.67" x 2.03" x 0.50"
Dimensions (L x W x H)	22 x 80 x 4 mm (Max)
Weight	6.6g

PRODUCT SPECIFICATIONS

NAND Components	3D Flash Memory
Interface	PCIe Gen 4 x4
Form Factor	M.2 2280

PERFORMANCE (MAX SEQUENTIAL READ/WRITE)

1TB	7,200 MB/s (Read), 6,500 MB/s (Write)
2TB	7,000 MB/s (Read), 5,900 MB/s (Write)

POWER

PCIe Power Management	APST, ASPM, L1.2

ENVIRONMENTAL

Operating Temperature	0°C ~ 70°C
Storage Temperature	-40°C ~ 85°C

RELIABILITY/SECURITY

Mean Time Between Failures (MTBF)	1,600,000 Hours
Error Correction Code	LDPC (Low Density Parity Check) ECC Algorithm
Product Health Monitoring	Self-Monitoring, Analysis and Reporting Technology (S.M.A.R.T)
Full End-to-End Data path protection	Supported
Encryption	TCG Pyrite

COMPATIBILITY

RAID 0	Supported
NVMe	NVMe 1.4
Performance Optimization	TRIM (requires OS support)

FOR MORE INFORMATION:

Contact your PNY Account Manager or email **GOPNY@PNY.COM** for more information.

PNY PRO