**AI Việt Nam bước vào kỷ nguyên mới: Đại học Công nghệ Thông Tin sở hữu siêu máy chủ AI NVIDIA DGX A100 đầu tiên tại Việt Nam**

**Trường đại học Công nghệ thông tin (Đại học Quốc gia TP.HCM) vừa tiếp nhận siêu máy chủ AI NVIDIA DGX A100 dựa trên kiến trúc Ampere.**

Siêu máy chủ NVIDIA DGX A100 là hệ thống phổ quát cho tất cả các cơ sở hạ tầng AI, từ phân tích (analytics), đào tạo (training) đến suy luận (inference).Nó đóng gói mật độ cao các tính toán với hiệu suất xử lý AI (artificial intelligence – trí tuệ nhân tạo) lên đến 5 petaFLOPS trong một khung máy 6U, thay thế một tháp cơ sở hạ tầng chỉ bằng một nền tảng cho mọi AI workload (công việc cần sử dụng trí tuệ nhân tạo).

Sử dụng nội bộ phục vụ cho công tác nghiên cứu khoa học

PGS.TS Nguyễn Hoàng Tú Anh – hiệu trưởng Trường đại học Công nghệ thông tin (Đại học Quốc gia TP.HCM) – cho hay đây là dòng máy chủ AI phục vụ cho chạy các ứng dụng, thuật toán AI.

* Siêu máy chủ NVIDIA DGX A100 nằm trong dự án chuyển đổi số và phục vụ chạy các mô hình deep leaning, máy học phục vụ nghiên cứu khoa học từ các nhóm nghiên cứu, phòng thí nghiệm và các đề tài nghiên cứu khoa học.

“*Trường chúng tôi là trường đại học đầu tiên ở Việt Nam sở hữu siêu máy chủ AI NVIDIA DGX A100. Nguồn kinh phí mua sắm máy chủ này từ đề án về chuyển đổi số và định hướng đại học số của Trường đại học Công nghệ thông tin do Đại học Quốc gia TP.HCM cấp từ ngân sách nhà nước”,* bà Anh cho biết thêm.

Được biết, để có siêu máy chủ phục vụ cho công tác nghiên cứu khoa học này, Trường đại học Công nghệ thông tin đã mất 9 tháng trao đổi để được NVIDIA đồng ý cung cấp và xuất khẩu khỏi Mỹ. Nhà trường phải ký cam kết sử dụng nội bộ phục vụ cho công tác nghiên cứu khoa học.



*Siêu máy chủ NVIDIA DGX A100 tại trường Đại học Công nghệ thông tin. Ảnh: M.K*



*Các nghiên cứu viên Trường đại học Công nghệ thông tin (Đại học Quốc gia TP.HCM) đang kiểm tra kết cấu và linh kiện siêu máy chủ AI NVIDIA DGX A100 vừa được nhập về trường – Ảnh: TRƯƠNG DŨNG*

Siêu máy chủ có bộ tăng tốc tiên tiến nhất thế giới

NVIDIA DGX A100 là hệ thống phổ quát cho tất cả các AI workload, cung cấp mật độ tính toán, hiệu suất và tính linh hoạt chưa từng có trong hệ thống AI 5 petaFLOPS đầu tiên trên thế giới.

Siêu máy chủ này có bộ tăng tốc tiên tiến nhất thế giới, đó là GPU NVIDIA A100 Tensor Core dựa trên kiến trúc Ampere.

NVIDIA DGXperts là một nhóm toàn cầu gồm hơn 14.000 chuyên gia và lập trình viên thông thạo AI, những người đã có nhiều kinh nghiệm trong thập kỷ qua để giúp bạn tối đa hóa giá trị của khoản đầu tư cho hệ thống DGX của mình.

DGX A100 là hệ thống AI đầu tiên trên thế giới được xây dựng trên GPU NVIDIA A100 Tensor Core. Tích hợp 8 GPU A100, hệ thống cung cấp khả năng tăng tốc chưa từng có và được tối ưu hóa hoàn toàn cho phần mềm NVIDIA CUDA-X™ và các lớp giải pháp end-to-end cho trung tâm dữ liệu NVIDIA.

Siêu máy chủ này có khả năng mở rộng cho trung tâm dữ liệu chưa từng có. Mở rộng trung tâm dữ liệu AI khi dữ liệu và bài toán bùng nổ là một yêu cầu không tránh khỏi. NVIDIA DGX A100 tích hợp Mellanox ConnectX-6 VPI HDR InfiniBand/Ethernet network adapters với 450GB/s băng thông hai chiều cao nhất.

Đây là một trong nhiều tính năng giúp DGX A100 trở thành khối xây dựng nền tảng cho các cụm AI lớn như NVIDIA DGX SuperPOD – một bản thiết kế Enterpise cho cơ sở hạ tầng AI có thể mở rộng.

***Trích dẫn từ tài liệu:***

https://adg.vn/tin-cong-nghe/ai-viet-nam-buoc-vao-ky-nguyen-moi-dai-hoc-cong-nghe-thong-so-huu-sieu-may-chu-ai-nvidia-dgx-a100-dau-tien-tai-viet-nam