

Компания Tesla представила систему Stress, предназначенную для выявления неисправных ядер в суперкомпьютере Dojo, состоящем из миллионов вычислительных блоков.

Обнаружение таких ошибок крайне важно, поскольку один незамеченный сбой может сорвать «многонедельную» сессию обучения искусственного интеллекта. Проблема тихого повреждения данных (SDC) широко распространена в процессорах с высокой плотностью ядер и длительными периодами работы.

Обучающие плитки Dojo с 8850 64-битными ядрами архитектуры RISC-V могут столкнуться с повреждением данных без явных признаков неисправности. Разработчики Stress добились того, что система назначает каждому ядру уникальную программу и позволяет обнаруживать неисправности во время работы без остановки процесса, связываясь с другими ядрами, что повышает эффективность обнаружения неисправностей в 10 раз.

Stress может выявлять неисправности на миллионах активных ядер, обычно ошибки обнаруживаются после обработки 1-100 ГБ тестовых инструкций, хотя в сложных случаях может потребоваться более 1 ТБ. Кроме того, Stress помогает выявлять ошибки в архитектуре и уязвимости программного обеспечения во время развёртывания.