GPU Memory Scanning: GPT 485B Model Training
(Megatron-LM, 8192 GPUs, A100-SXM-80GB, FP16)