GPU Memory Scanning: 485B Model Training Using Megatron-LM
(8192 GPUs, A800-SXM4-80GB, FP16)