seq4096-hd128-16h-32b | Q=4096, K=4096, head_dim=128, heads=16, batch=32

Kernel	Avg Forward Time (ms)	Relative to Triton Flash
scaled_dot_prod_attention	ERROR	N/A
vectorized_torch	ERROR	N/A
vectorized_torch_compiled	ERROR	N/A
flash_attention_triton	**493.08**	**1.00x**