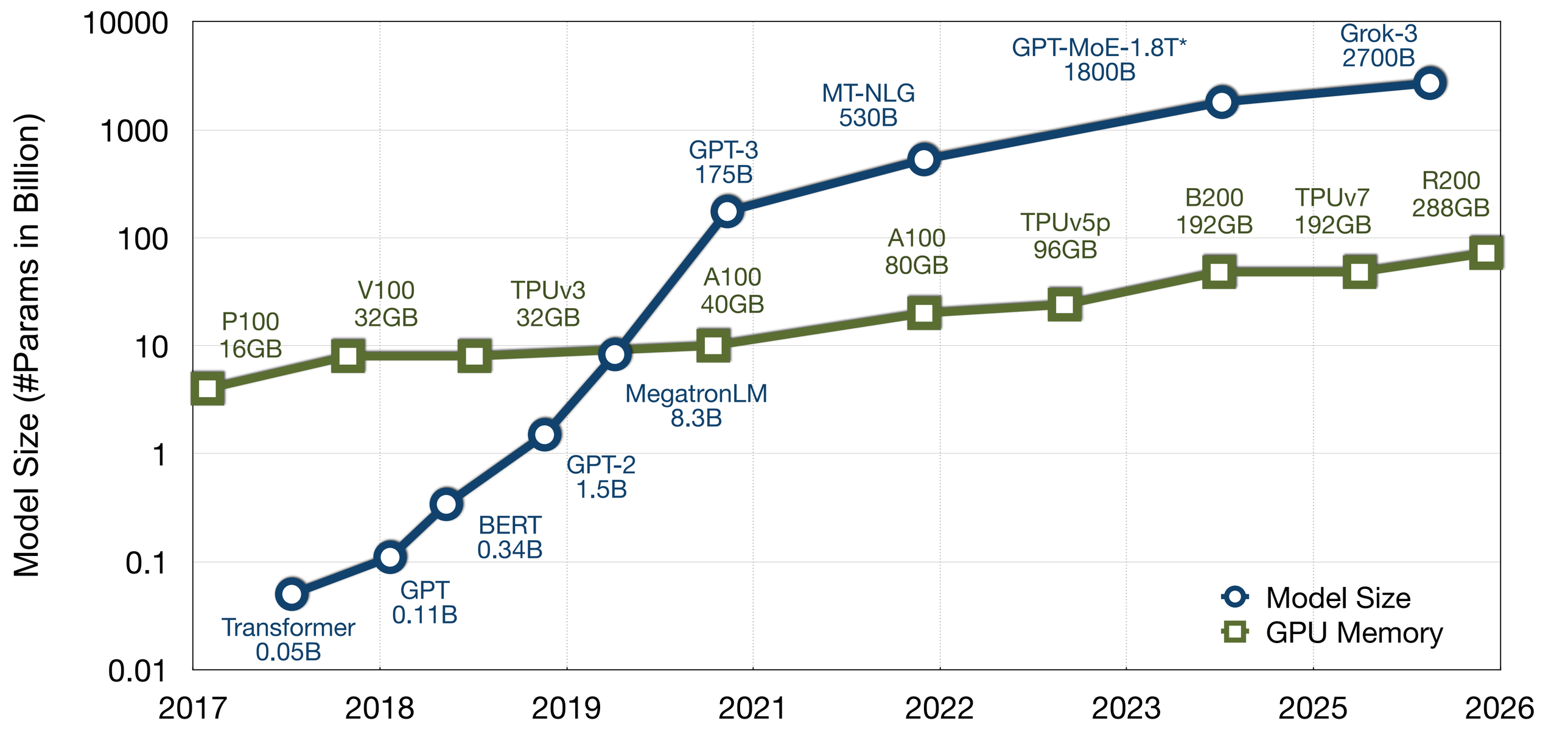

Model Size (#Params in Billion)
10000
1000
100
10
1
0.1
0.01
2017
2018
2019
2021
2022
2023
2025
2026
Transformer
0.05B
GPT
0.11B
BERT
0.34B
GPT-2
1.5B
MegatronLM
8.3B
GPT-3
175B
MT-NLG
530B
GPT-MoE-1.8T*
1800B
Grok-3
2700B
P100
16GB
V100
32GB
TPUv3
32GB
A100
40GB
A100
80GB
TPUv5p
96GB
B200
192GB
TPUv7
192GB
R200
288GB
Model Size
GPU Memory