Training Speedup on GPU cluster with 1Gbps Ethernet

Legend:
- AlexNet Baseline
- AlexNet TernGrad
- AlexNet Deep Gradient Compression
- DeepSpeech Baseline
- DeepSpeech TernGrad
- DeepSpeech Deep Gradient Compression

Y-axis: Speed up (0 to 60)

X-axis: # of training nodes (1, 2, 4, 8, 16, 32, 64)