

PowerElecBench Evaluation Results Summary

Model	L1 (150)	L2 (150)	L3 (100)	L4 (100)	Total	Cost	Latency
Grok 4.1 Fast	92 (61.3%)	96 (64.0%)	54 (54.0%)	21 (21.0%)	263 (52.6%)	\$0.12	~4s
GPT-4o	91 (60.7%)	88 (58.7%)	57 (57.0%)	9 (9.0%)	245 (49.0%)	\$3.00	~12s