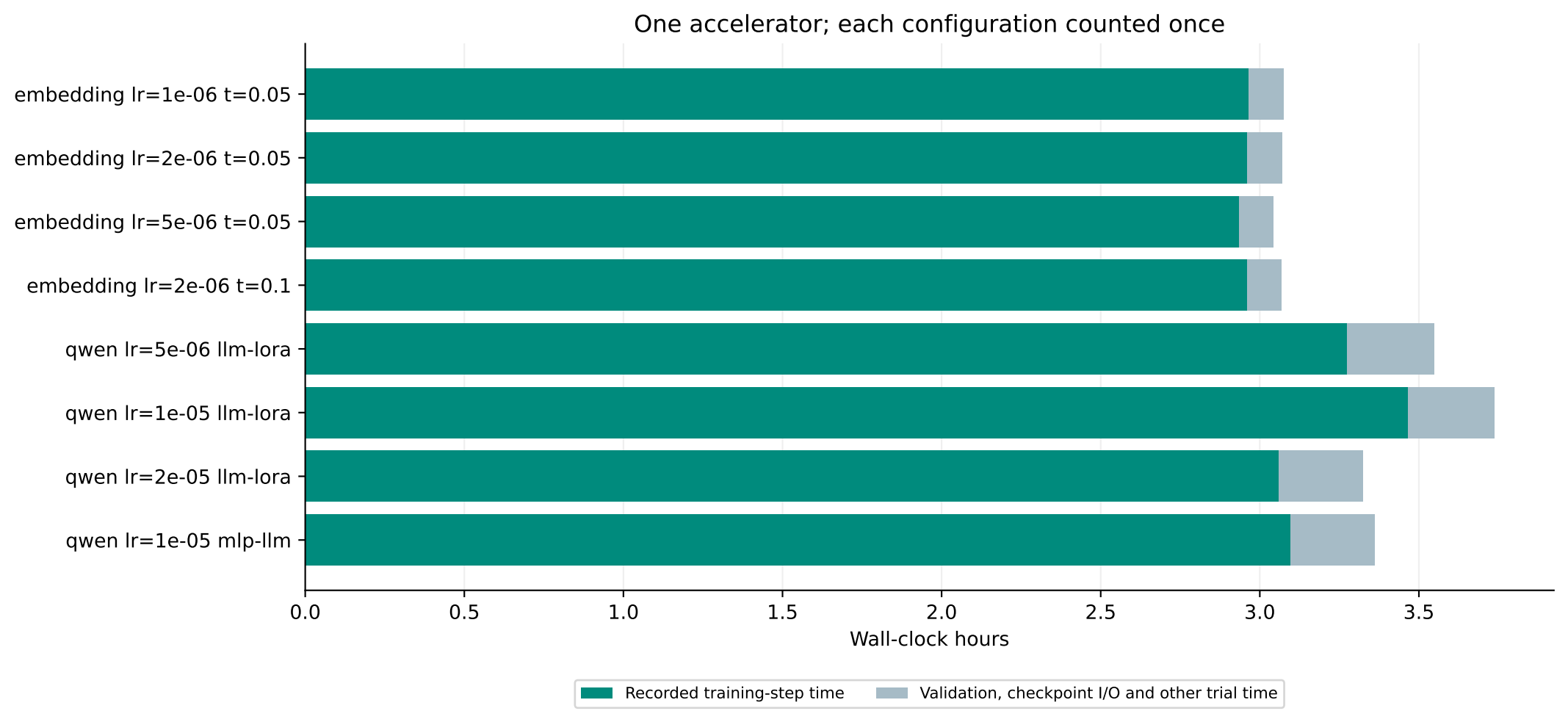

One accelerator; each configuration counted once
embedding lr=1e-06 t=0.05
embedding lr=2e-06 t=0.05
embedding lr=5e-06 t=0.05
embedding lr=2e-06 t=0.1
qwen lr=5e-06 llm-lora
qwen lr=1e-05 llm-lora
qwen lr=2e-05 llm-lora
qwen lr=1e-05 mlp-llm
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Wall-clock hours
Recorded training-step time
Validation, checkpoint I/O and other trial time