Memory Configuration Performance Impact

Throughput (tokens/sec) vs **Memory Configuration**

- No optimizations: 45,935 (+0.0%)
- Pool + Unified (128KB): 72,946 (+58.8%)
- Pool + Unified (256KB): 39,913 (-13.1%)
- Pool only: 41,022 (-10.7%)
- Unified (128KB) only: 48,963 (+6.6%)