vLLM Reliability Matrix: Success Rate by Configuration
(Llama-3.1-8B on A100 40GB)