RAG Strategy: Accuracy vs Latency