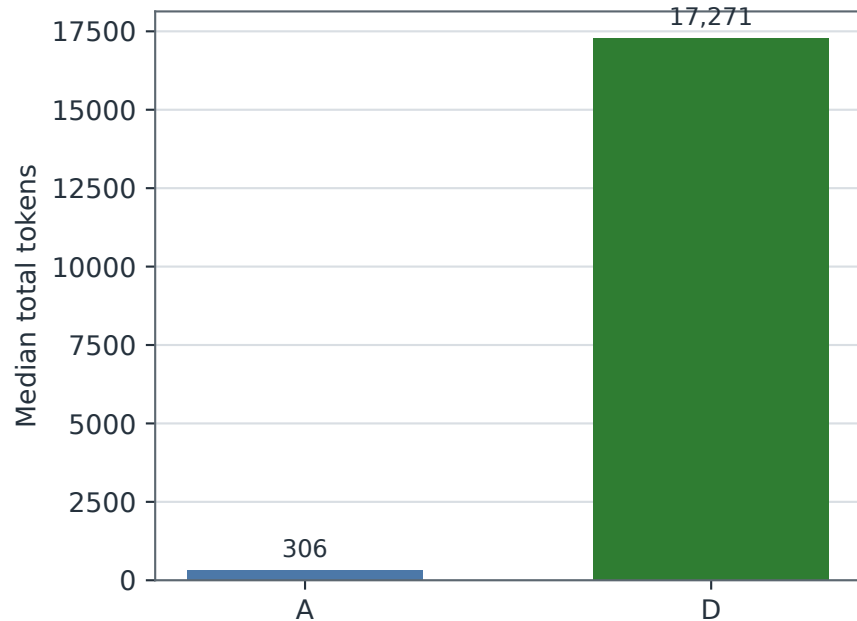
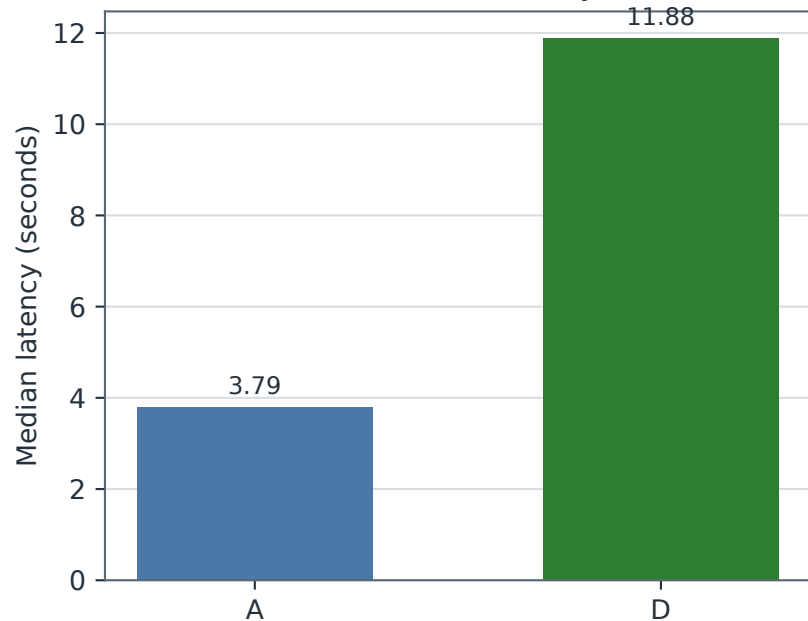


Operational trade-off of the guardrailed retrieval bundle

Token burden



End-to-end latency



Both paired differences were significant by Wilcoxon signed-rank test ($p < 0.001$).