

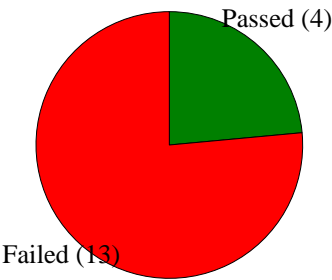
Router Agent Performance Test Suite - Performance Report

Agent: router
Version: 1.0
Description: Comprehensive performance and accuracy tests for the router agent to validate routing decisions between direct_response and research_workflow
Generated: 2025-08-23 02:11:48

Executive Summary

Metric	Value
Total Tests	17
Passed	4
Failed	13
Success Rate	23.53%
Mean Latency	0.21s
Median Latency	0.00s
Total Tokens	12,158
Total Cost	\$0.0037

Test Results Distribution



Detailed Test Results

Scenario	Result	Latency	Tokens	Cost
Simple Greeting	✓	0.93s	2,419	\$0.0007
Morning Greeting	✓	0.60s	2,425	\$0.0007
Thank You Message	✓	0.82s	2,427	\$0.0007
Capability Question	✓	0.55s	2,431	\$0.0007
User Clarification	✗	0.68s	2,456	\$0.0007
RBC Efficiency Ratio Request	✗	0.00s	0	\$0.0000
Bank Comparison Request	✗	0.00s	0	\$0.0000
Trend Analysis Request	✗	0.00s	0	\$0.0000
Financial Statement Request	✗	0.00s	0	\$0.0000
Benchmark Data Request	✗	0.00s	0	\$0.0000
Vague Numbers Request	✗	0.00s	0	\$0.0000
Greeting with Data Request	✗	0.00s	0	\$0.0000
Contextual Follow-up	✗	0.00s	0	\$0.0000
Calculation Request	✗	0.00s	0	\$0.0000
Report Generation Request	✗	0.00s	0	\$0.0000
Long Conversation Context	✗	0.00s	0	\$0.0000
Complex Multi-part Query	✗	0.00s	0	\$0.0000

Failed Test Details

User Clarification

Scenario ID: clarification_response

Expected: direct_response

Actual: research_workflow

Error: Expected route 'direct_response', got 'research_workflow'

RBC Efficiency Ratio Request

Scenario ID: efficiency_ratio_request

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Bank Comparison Request

Scenario ID: comparison_request

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Trend Analysis Request

Scenario ID: trend_analysis

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Financial Statement Request

Scenario ID: financial_statement

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Benchmark Data Request

Scenario ID: benchmark_request

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Vague Numbers Request

Scenario ID: vague_numbers

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Greeting with Data Request

Scenario ID: mixed_greeting_data

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Contextual Follow-up

Scenario ID: contextual_followup

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Calculation Request

Scenario ID: calculation_request

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Report Generation Request

Scenario ID: report_generation

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Long Conversation Context

Scenario ID: long_conversation

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe

Complex Multi-part Query

Scenario ID: complex_query

Expected: {'route': 'research_workflow'}

Actual: None

Error: Unexpected error: [Errno 32] Broken pipe