

ONE REPORTED FAILURE. A CHECKABLE RESULT.

# The tool returned. The agent stayed silent.

We replayed a reported Pipecat timing bug in a real local pipeline. The tool returned data, but the original revision didn't produce the next answer within our observation window.

| ORIGINAL               | PROPOSED PATCH         |
|------------------------|------------------------|
| <b>6 / 24</b>          | <b>24 / 24</b>         |
| selected checks passed | selected checks passed |

Eight predefined sequences, three repeats per revision. Consistency checks, not independent samples or a production reliability score.

## We checked the answer, not just the tool.

The failing sequence resumes speech before tool commitment and stops it while input is muted. The result arrives, but speaking state remains stuck and the next inference is held.

We also tried earlier and later speech stops, a slow tool, a handled tool error, no mute and two speech bursts. The proposed correction passed all eight sequences. Ordinary controls kept working.

The patch can recover through the watchdog; recovery isn't always immediate. This test doesn't establish recovery of muted words or cancellation of stale inference.

Real pipeline, aggregators, turn controller and function dispatch. Synthetic speech events and scripted model/tool responses. No STT, TTS, telephone network or customer deployment was tested.

Original report: Shehroze Talat, Pipecat #6201. Proposed fix: Cross2pro, PR #6232. Our work is independent verification, not the original discovery or patch. The PR was open when checked.

[Read the case, traces and reproduction commands ↗](#)

BEFORE THE CLIENT HANDOVER

# Bring one failure. Keep the test.

One difficult call flow. An independent check before the client handover.

## What you get

One agreed workflow, a runnable reproduction where possible, and the before/after evidence. We check nearby behavior that must keep working, then hand over the commands, test inputs and limitations.

STARTING PILOT OFFER

# €1,500

Before applicable taxes. Final scope agreed in writing.

One workflow. Up to five working days once access is ready. Maximum 15 engineering hours. We agree on the acceptance criteria and any model-call budget before starting.

If the failure doesn't reproduce within budget, you still get the tested cases, results and gaps. We don't guarantee a defect, a production fix or exhaustive coverage.

## Start without sending secrets.

Send your stack, what happened and what should have happened. We'll agree on synthetic data or approved staging before testing.

No production changes, live security testing, training or private customer records by default. No security certification.

IGOR • PRAGUE, CZECH REPUBLIC

[laplanddream@gmail.com](mailto:laplanddream@gmail.com)[ibackstrom.github.io/faultline](https://ibackstrom.github.io/faultline)