TESTINGMIND

Smarter Testing for AI Generated Code

New Strategies for an Agentic AI World

Matthew LeRay
CTO/Co-Founder, Speedscale

Adapting QA to AI Generated Code

01 Challenges with AI Coding Assistants

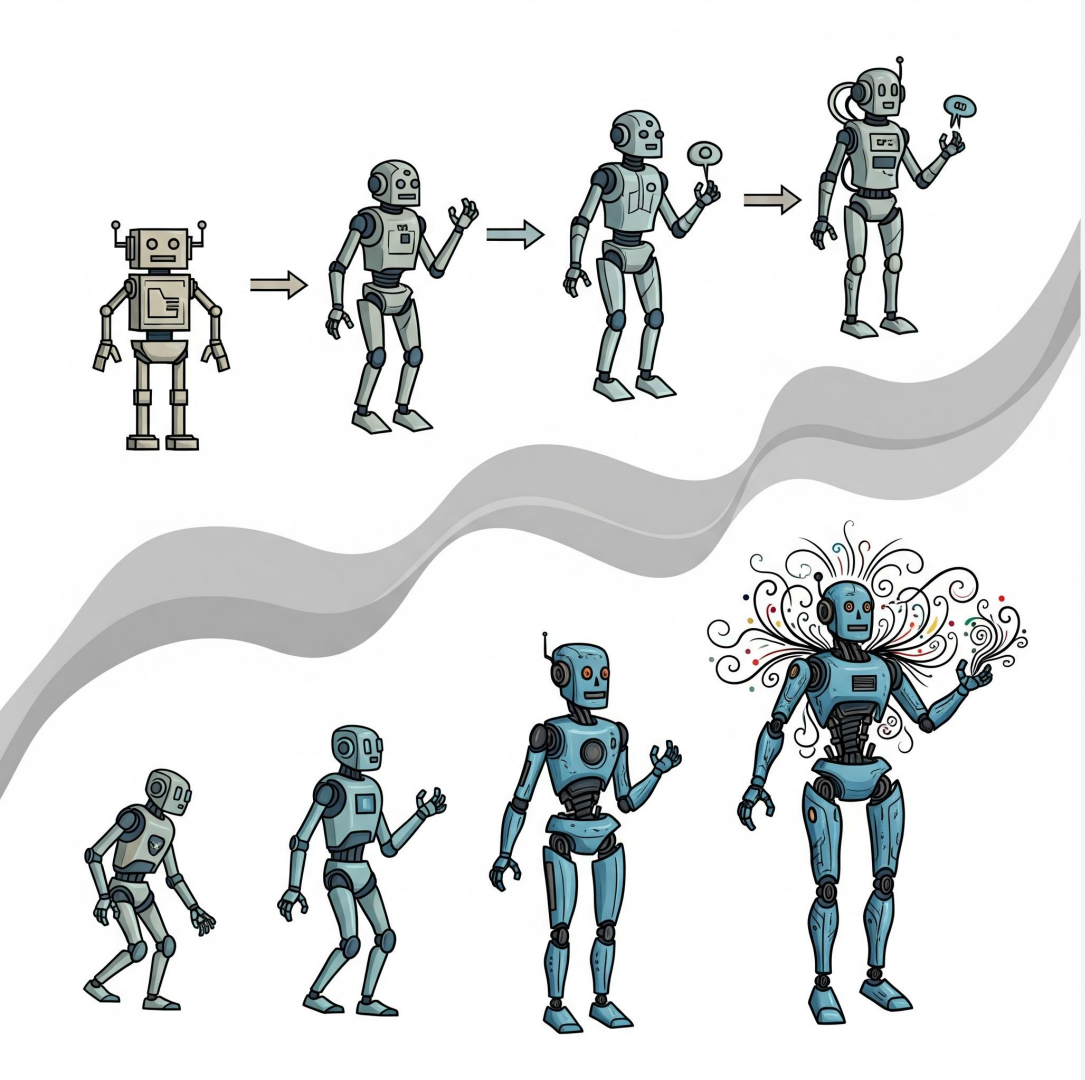
02 Impact on QA

03 Case Studies & How to Improve

04 On the Horizon

AI Coding

Solid ROI... but with trust issues



Agentic AI Stages

Rule-based Workflow: AI follows predefined rules.

Basic (Executor): AI provides simple, direct answers.

Use of Tools (Actor): AI utilizes tools to complete tasks.

Observe, Plan, Act (Operator): AI plans and executes actions.

Fully Autonomous (Explorer): AI operates independently.

Fully Creative (Inventor): AI generates novel solutions.

source:vellum

AI Coding != Business Process Automation

With AI, RPA robots can emulate a wide range of the decision making and problem solving that people do—which expands the range of work that you can automate. Here is where we've prioritized the use of AI across the UiPath Business Automation Platform:

Discover Opportunities

We use AI to discover – uncovering process improvement and automation opportunities from the work done between systems, people and communications

Processes
Model enterprise processes & deduce organizational processes

Tasks
Identify automatically or define through SMEs the ideal process

Communications
Gather insights from business conversations like emails and tickets

Automate More

We use AI to automate more – adding new skills to our robots allowing humans to be more creative and innovative

- Communications
- Interfaces
- Documents
- Applications

Supercharge Productivity

We use AI to accelerate common tasks – by bringing together automation and industry-leading generative AI functionality to the desktop

Business + Developers + IT

Agentic AI aka **Robotic Process Automation.**

Enterprise IT hosts Agentic platforms for business analysts to build agents on.

n8n

HIGHLY ADOPTED ENTERPRISE

AI Coding Tool Adoption

- Prototyping tools like Windsurf/Lovable are popular for **“vibe coding”**
- Enterprises use:
  - Copilot: Microsoft-focused or security-oriented
  - Cursor: IDE-based
  - Claude: CLI or Agent-based
- Cursor had an early lead but Claude is accelerating

OpenAI

bolt.new

Lovable

source: StackOverflow

AI Adoption

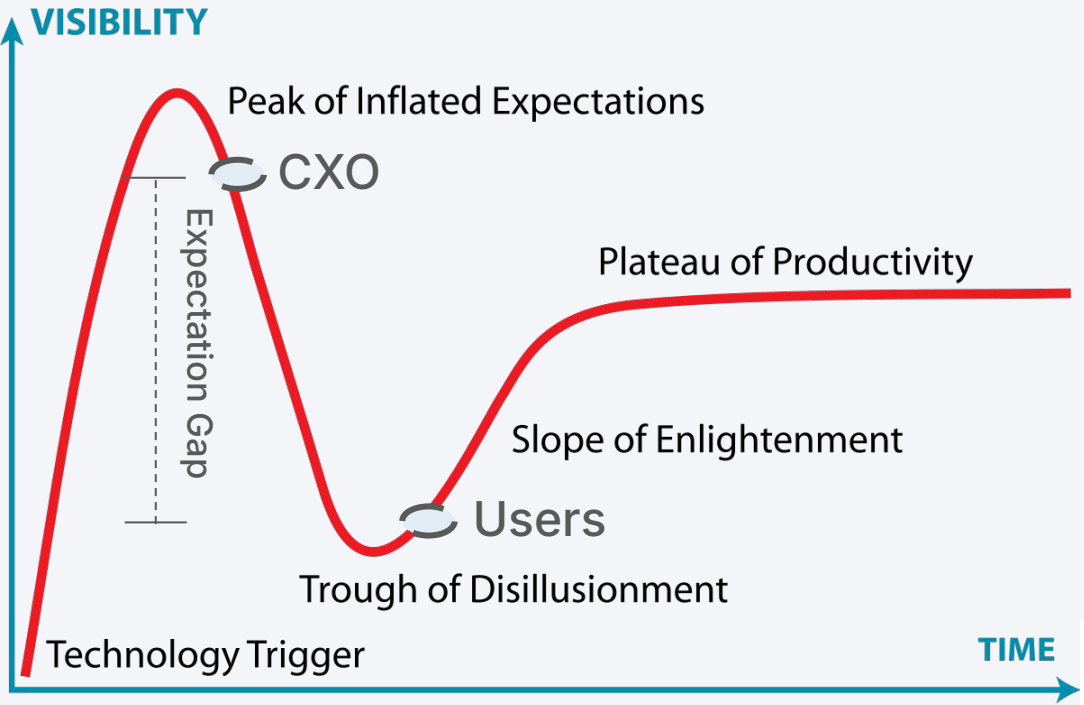
“84% of respondents are using or planning to use AI tools in their development process, an increase over last year (76%). This year we can see 51% of professional developers use AI tools daily.”

source: StackOverflow

… but confidence in AI coding tools lags

“More developers actively distrust the accuracy of AI tools (46%) than trust it (33%), and only a fraction (3%) report "highly trusting" the output. Experienced developers are the most cautious, with the lowest "highly trust" rate (2.6%) and the highest "highly distrust" rate (20%), indicating a widespread need for human verification for those in roles with accountability.”

AI ROI is Unclear



- **Users** see value but with limitations
- **Executives** are promised sizable cost savings
- Plateau of Productivity is approaching as these converge

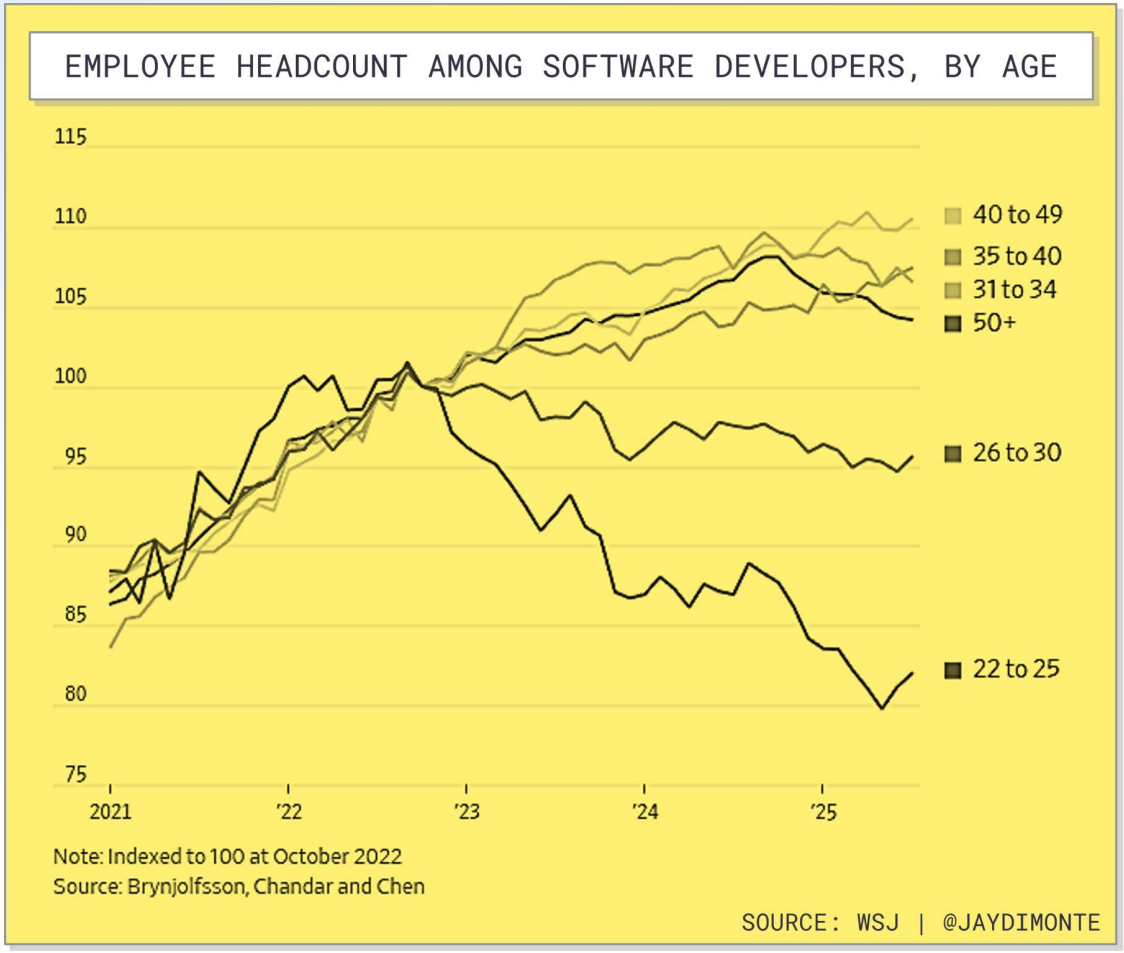
Astasia Myers · 2nd
General Partner at Felicis
1h ·

2/ "Vibe ARR" is everywhere
Remember Jasper? hyper growth, but the model companies ate them alive. We're seeing this again. Companies growing really quickly on point-in-time technology that isn't defensible

source: MIT

Developer Impact

- **Executives** pushing use of AI in known areas like software engineering
- HR slowing hiring of new grads
- Engineering Leaders on the hook to deliver more with less team members



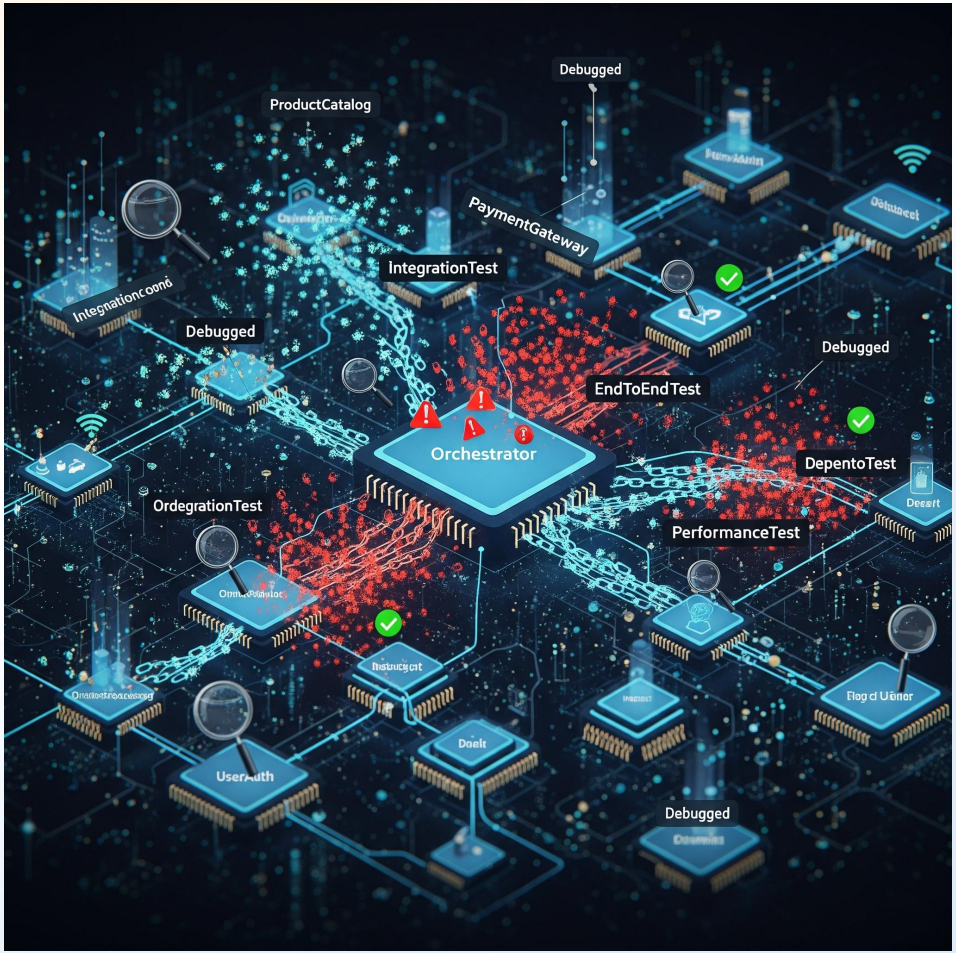
Source: [Jackie DiMonte]

QA Landscape

Agile, Microservice and now... AI

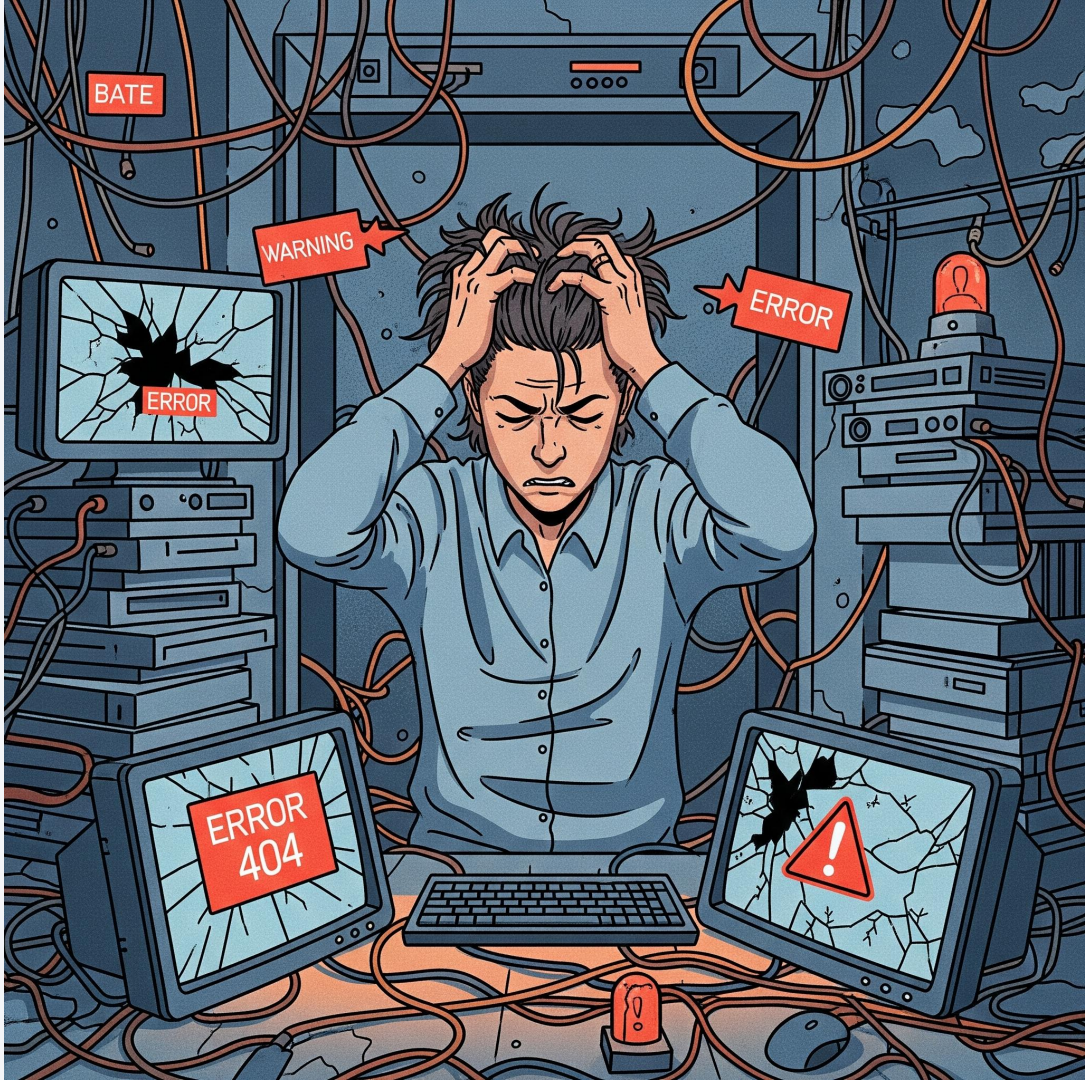
Microservices

- Focus shifts to inter-service interactions
- Integration testing is more important but also less effective



Environment Challenges

- Test data synchronization is a major issue
- Multiple users compete for limited environments
- 3rd Party Systems can’t be easily mocked or simulated



AI and Determinism

AI coding agents produce stochastic results.

QA relies on deterministic outcomes.

Summary of Common Challenges

AI Produces Lots of Broken Code

QA Test Environments are not Microservices-Enabled

AI is Stochastic, Testing is Deterministic

Testing is too far “right”

Solutions

Process, Tools, Expectations

Process: Treat Tests and Environments as Cattle, not Pets

- Use LLM tools to generate, manage and analyze test flows
  - Fuzzing
  - Edge cases
- Replicate Environments, not mocks
  - Ephemeral Environments
  - Developer Portals
  - Simulation Engines

Video: YouTube

Add an LLM-Powered Code Review CI Step

Claude

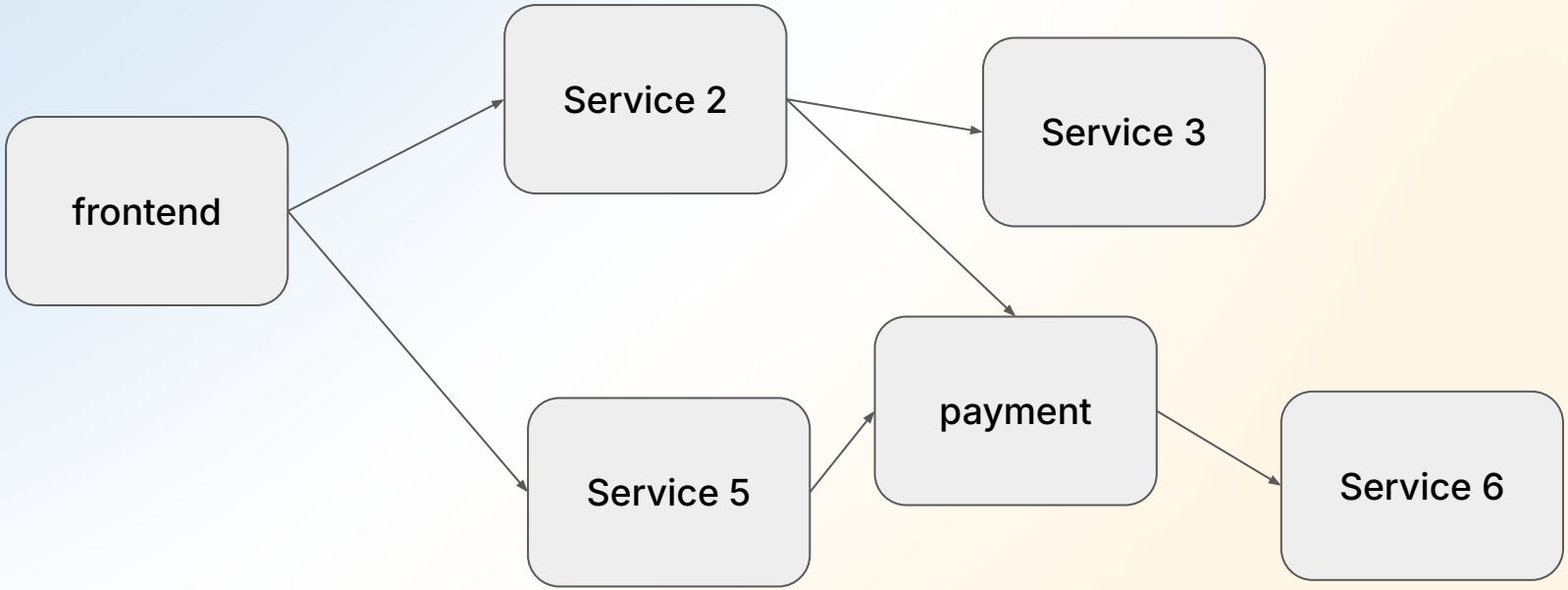
CodeRabbit

Tricentis

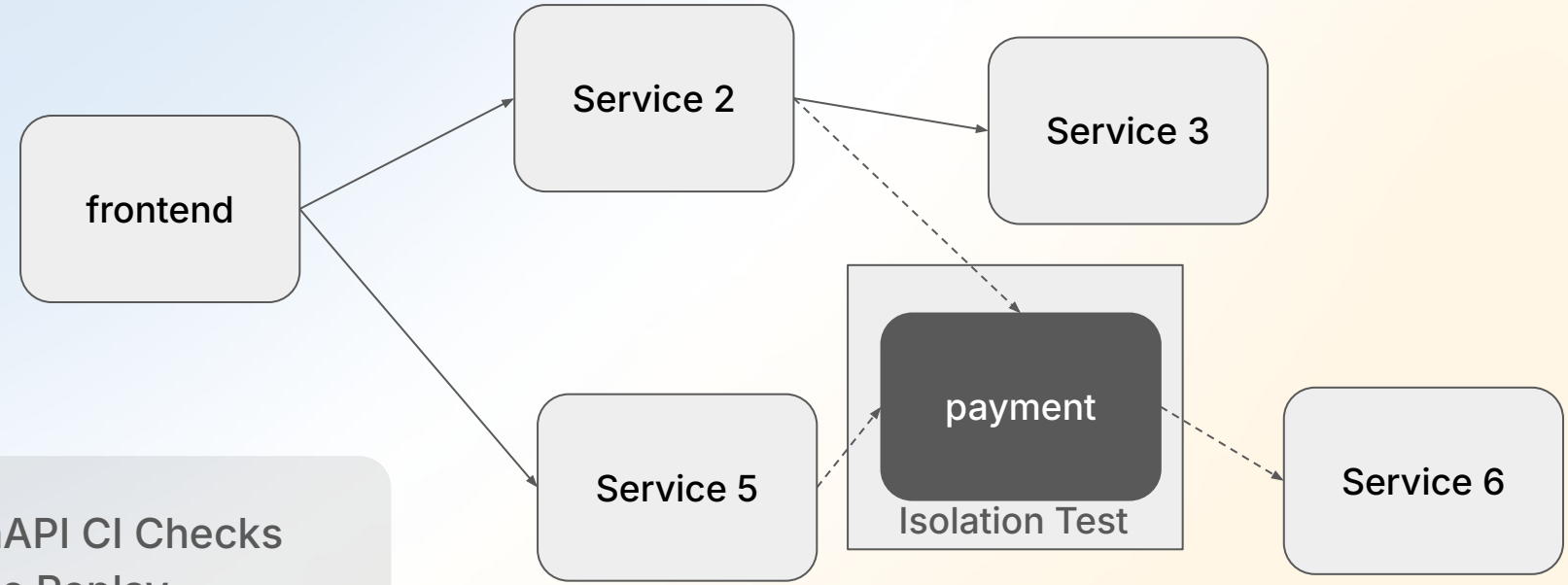
"We evaluate Code Researcher on kBenchSyz [23], a benchmark of Linux kernel crashes, and show that it significantly outperforms strong baselines, achieving a crash-resolution rate of 58%, compared to 37.5% by SWE-agent [40]"

Research Credit: Microsoft

Replace (some) Integration Tests with Isolation Tests

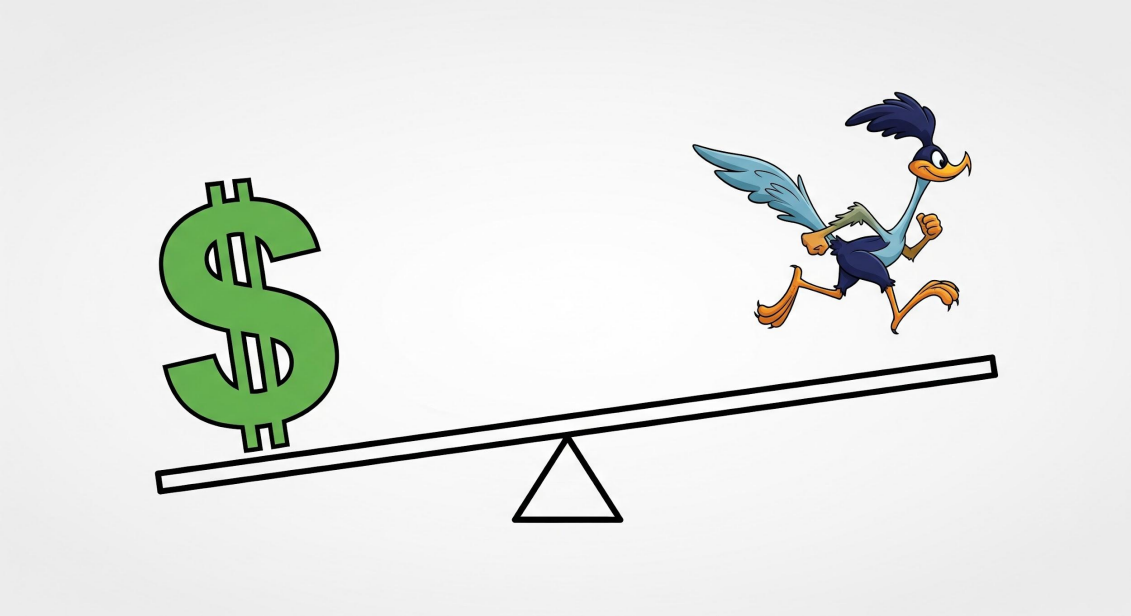


Replace (some) Integration Tests with Isolation Tests



- OpenAPI CI Checks
- Traffic Replay
- Simulation Environments

Test for FinOps, not stress or performance



“Anything is Possible (Given Enough Time and Money)” - My Contractor

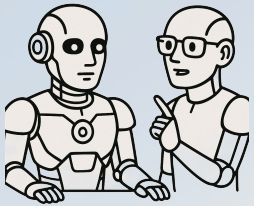
Tactical Tips and Tricks

- All artifacts in **Markdown**
- Provide centralized testing services via MCP
- Advocate for QA as part of the AI “inner loop”
- Automate everything and build agents for some tasks
  - Summarization
  - Test case generation
  - DevOps
  - Project Management

On the Horizon

On The Horizon

Pair Programmer

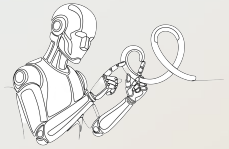


MCPs that can answer deeper questions about the codebase, specs or real world function.

Agentic AI for QA

AI Agent capable of testing new builds unattended.

Closed-loop Ticket Resolution



AIs finding bugs, opening tickets and resolving them without human interaction

Thank you.

Matthew LeRay, CTO/Co-founder
matt@speedscale.com