



**ENTERPRISE  
DEVOPS**  
T E C H C O N

## The Cloud Native Stack & You

Coté, October 2nd, 2023



Hello, I am Coté.

I work at **vmware**<sup>®</sup>

More: <https://cote.io>



But first..

OP

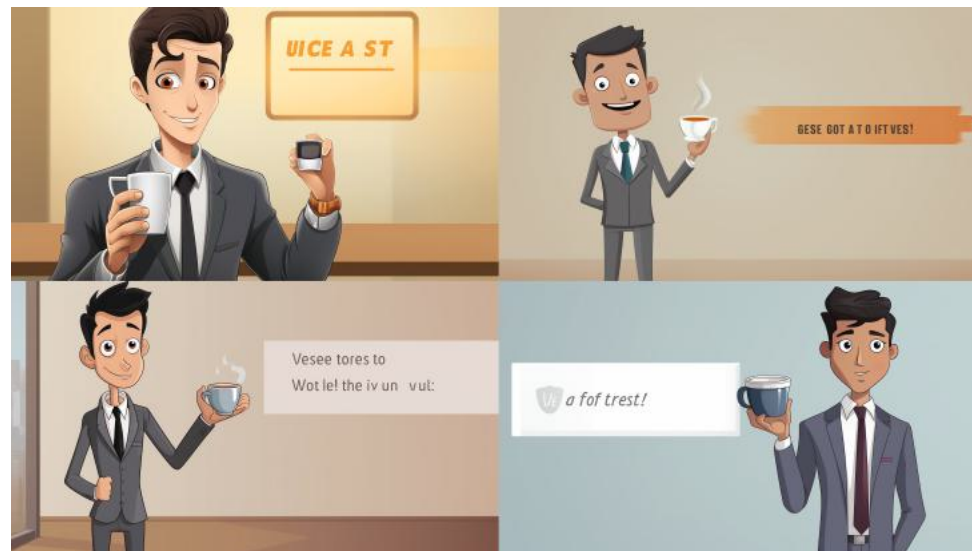
What is a good image generation prompt for "a presentation slide that connotes the thought 'you want me to talk about AI, which I kind of don't care about, but, here, I'll talk to you about it.'"



Ah, a slide that exudes a certain nonchalance while still fulfilling the task at hand—quite the rhetorical challenge. How about this for an image generation prompt:

"A minimalist presentation slide with the title 'AI: Because You Asked For It.' The background subtly alternates between varying shades of gray, symbolizing indifference. On one corner, a cartoon avatar of the presenter is depicted with one hand in a pocket and the other holding a coffee cup. The avatar has a speech bubble with the words 'Let's get this over with.'"

This should capture the essence of reluctant participation while still maintaining a sense of professionalism. What do you think?

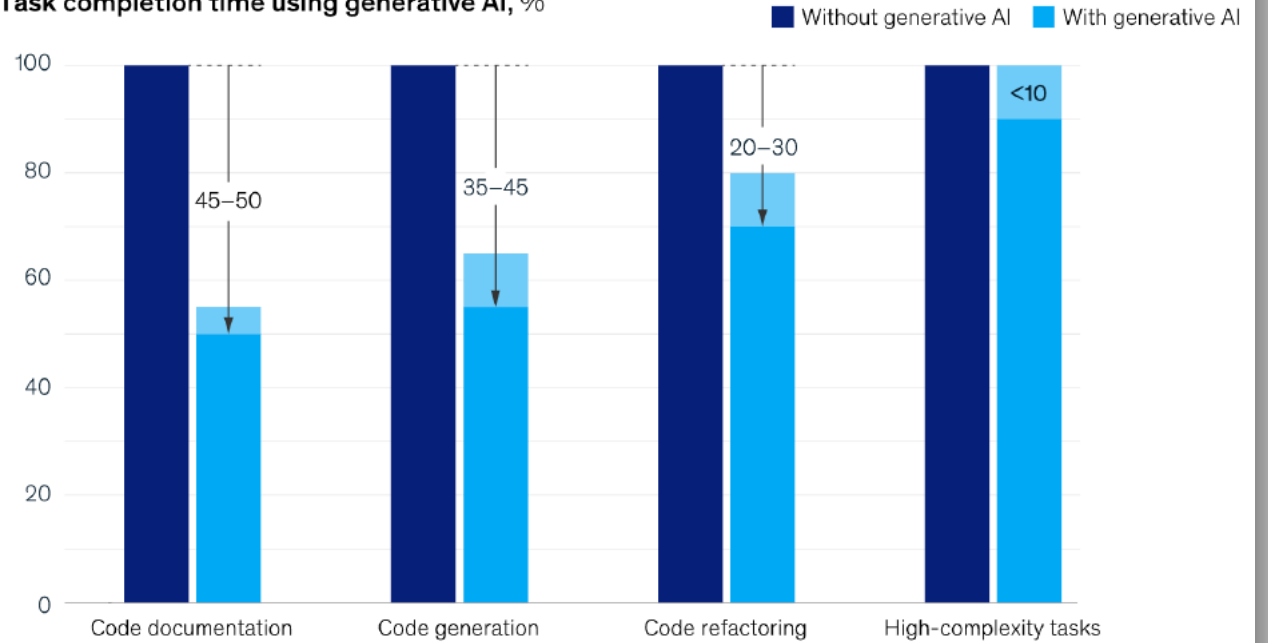


# You should check it out! It's great!

- 56% use AI frequently or occasionally.
- 68% for written content
- 60% for brainstorming ideas
- 50% for background research
- 19% for analysis & marketing forecasts
- 11% for generating computer code
- 7% for image recognition & generation

## Generative AI can increase developer speed, but less so for complex tasks.

Task completion time using generative AI, %



McKinsey & Company

Sources: [summary of Conference Board survey](#), Sep 2023; ["Unleashing developer productivity with generative AI,"](#) McKinsey, July, 2023. For general UI usage in the US, see ["Survey finds relatively few Americans actually use \(or fear\) ChatGPT,"](#) Deven Coldeway, TechCrunch, Aug, 2023.

OP

What is the bill gates quote about people over estimate a technology in the first year and underestimate it ten years?



The quote from Bill Gates goes something like this: "We always overestimate the change that will occur in the next two years and underestimate the change that will occur in the next ten."

It's a rather clever insight that speaks to the human tendency to be overly optimistic about the immediate future of technology, yet somewhat myopic about its long-term implications and potential transformations. It's a cautionary reminder to balance hype with a longer view, something that's particularly relevant in the fast-paced world of tech. Would you agree?

### Intelligent Assist

What needs my attention today?

I need to cut some cost, what can I do?

Show me Kubernetes Clusters that need my attention related to acme-fitness app

Can you show me my apps I should be concerned about?

Intelligent Assist

Here are 3 apps with cost, security, and policy violations:

- acme-fitness**  
VMW.Aria.BusinessApplication  
Version: 2.3  
Format: Carvel
- acme-nutrition**  
VMW.Aria.BusinessApplication  
Version: 2.3  
Format: Carvel
- acme-coaching**  
VMW.Aria.BusinessApplication  
Version: 2.3  
Format: Carvel

Take me to acme-fitness

Take me to acme-nutrition

acme-coaching

Text to chat.

### VMware Tanzu Hub

## Application Topology

Intelligent Assist

What do you think?

If you're ready, I can help you implement the recommended change. Do you want to do that?

I'm ready. Make the change.

Intelligent Assist

Perfect. One moment...

Done! **ACME-EU-FE-1** size changed. You'll save **\$680.00/mo (67%)** moving forward!

**ACME-EU-FE-1**  
AWS.EC2.Instance

Title: aws s q3-finance-03f-flg-app...  
ID: i-071468e-5eac3d55e  
Organization: Production  
Node: r5\_2xlarge  
CPU: 6 Cores  
Memory: 4 GB  
Disk: 11 GB

Take me back to acme-fitness

Intelligent Assist

No problem. I'm taking you back to where we started...

Here you are. Notice the critical violation has been fixed. Happy I could help!

Text to chat.

acme-fitness  
Hosted on AWS and Azure

Namespaces: 2

EKS Clusters: 4

Deployments: 3

Load Balancers: 2

IAM Role: 1

VIEW SERVICE MAP

SETTINGS

acme-fitness  
VMW.Aria.BusinessApplication

Status: Attached  
Kubernetes Version: v1.23.13-eks-...  
Cloud Account: aws.ec2.Instance  
Number of Manager Nodes: 1  
Number of Worker Nodes: 1

CLOSE EDIT DELETE

### Violations

Critical: 0 | High: 0 | Medium: 2 | Low: 1

Top 3

- This is a Security violations (Medium)
- This is a Policy violations (Medium)
- This is a Policy violations (Medium)

VIEW ALL VIOLATIONS

### Application cost

Projected cost (USD): **\$3,200 K**  
Potential Savings (USD): **\$64 K**  
▲ 2% over last month

Monthly cost by provider

Provider	Cost (USD)
AWS	\$2,261 M
Azure	\$975,670

Top Resizing Recommendations

Instance	Recommendation	Potential Savings
ACME-EU-FE-1	Resize to r5.2xlarge	\$230

Yup. AI. It's pretty cool\*

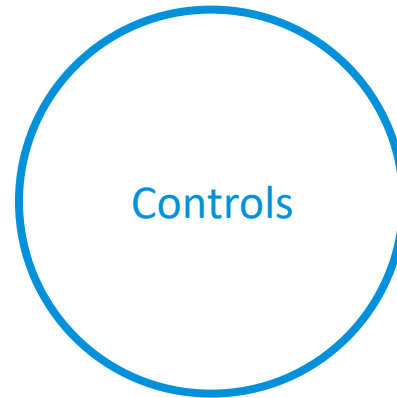
\* If you're *my* bank, please be safe.



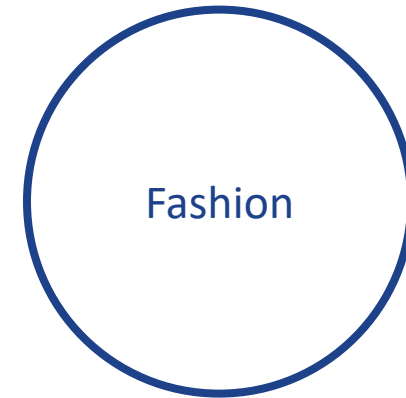
# Why do people want cloud native?



“Speed”  
Decoupling  
Eventual Toil Reduction



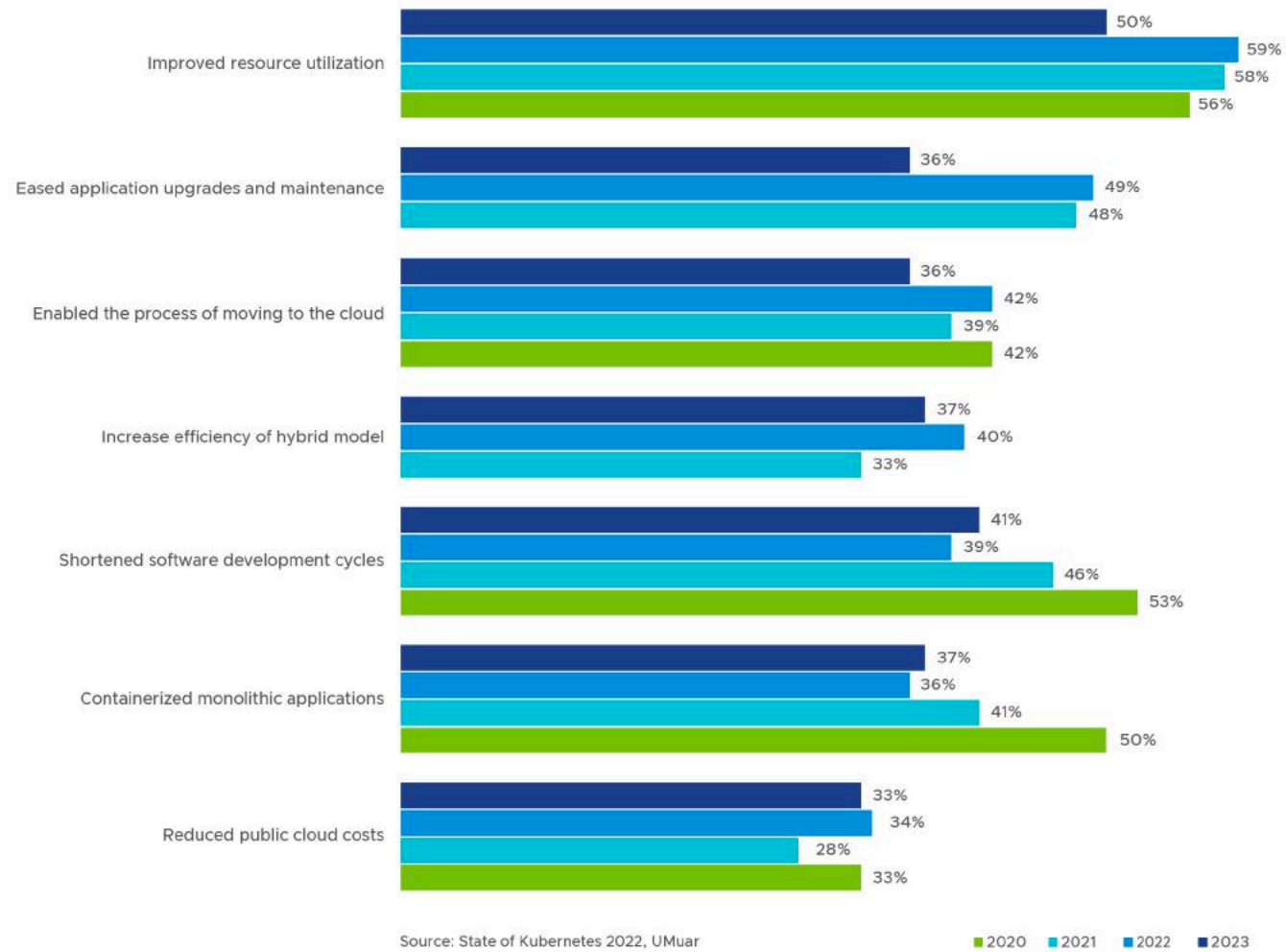
Security  
Governance  
Multi-cloud



Be like Google  
Kubernetes!  
Generational

### What benefits has your organization realized from operating Kubernetes?

(Choose all that apply)



Source: [State of Kubernetes 2023](#), VMware – analysis by Coté.

Q

Wait, so by “cloud native,” don’t you just mean “Kubernetes”?

A: Yes. No. Yes. No, *definitely* no.

*Maybe?*

Ugh.

(Cloud Foundry forever!)

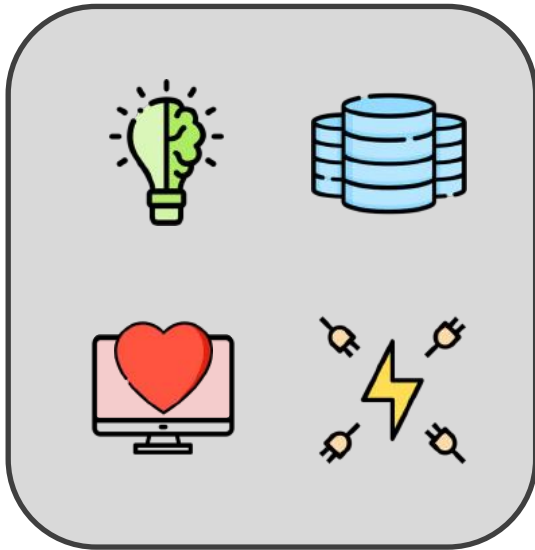
“ Cloud native technologies empower organizations to build and run scalable applications in modern, dynamic environments such as public, private, and hybrid clouds. Containers, service meshes, microservices, immutable infrastructure, and declarative APIs exemplify this approach.

These techniques enable loosely coupled systems that are resilient, manageable, and observable. Combined with robust automation, they allow engineers to make high-impact changes frequently and predictably with minimal toil.”

CNCF, June, 2018

# From monoliths to cloud native apps

## Monolith

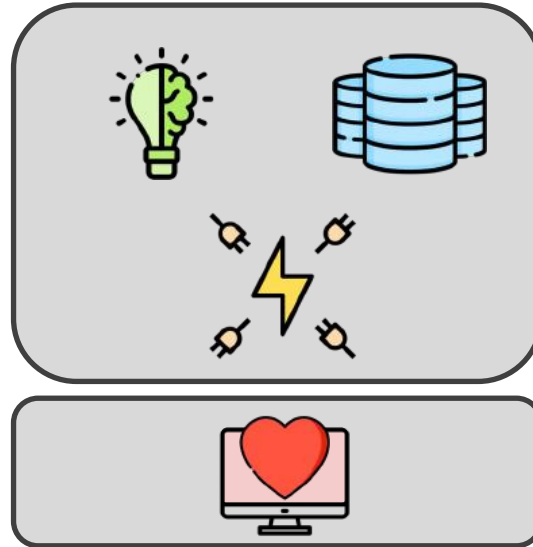


All components in one runtime.

All processing takes place in that runtime.

No network reliance.

## Client Server



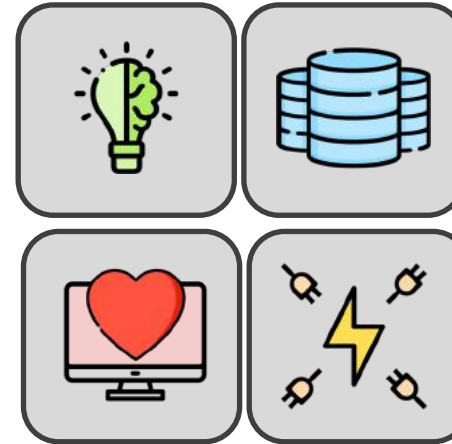
Access components run in client runtime.

All other app components run in one runtime.

Processing takes place in client and in app component runtime.

Client to Server communications over network.

## N-Tier

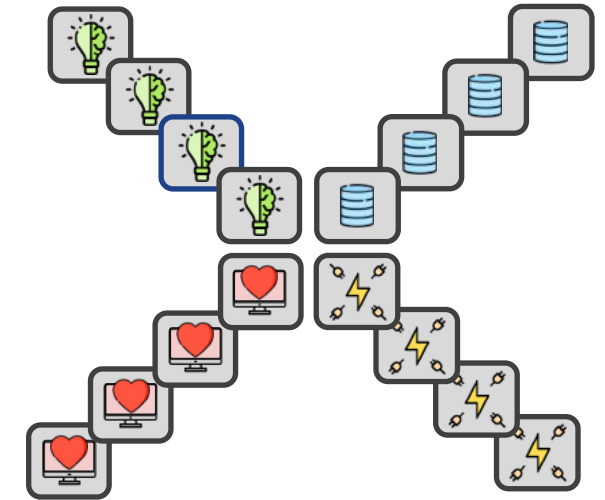


Each component in separate runtime.

Processing takes place in each runtime.

Components communicate over a network.

## Cloud Native



Each component broken down into services all running in own runtime.

Processing takes place in individual runtime.

High network dependency for all components to communicate over network.

Source: Marc Zottner, 2023.

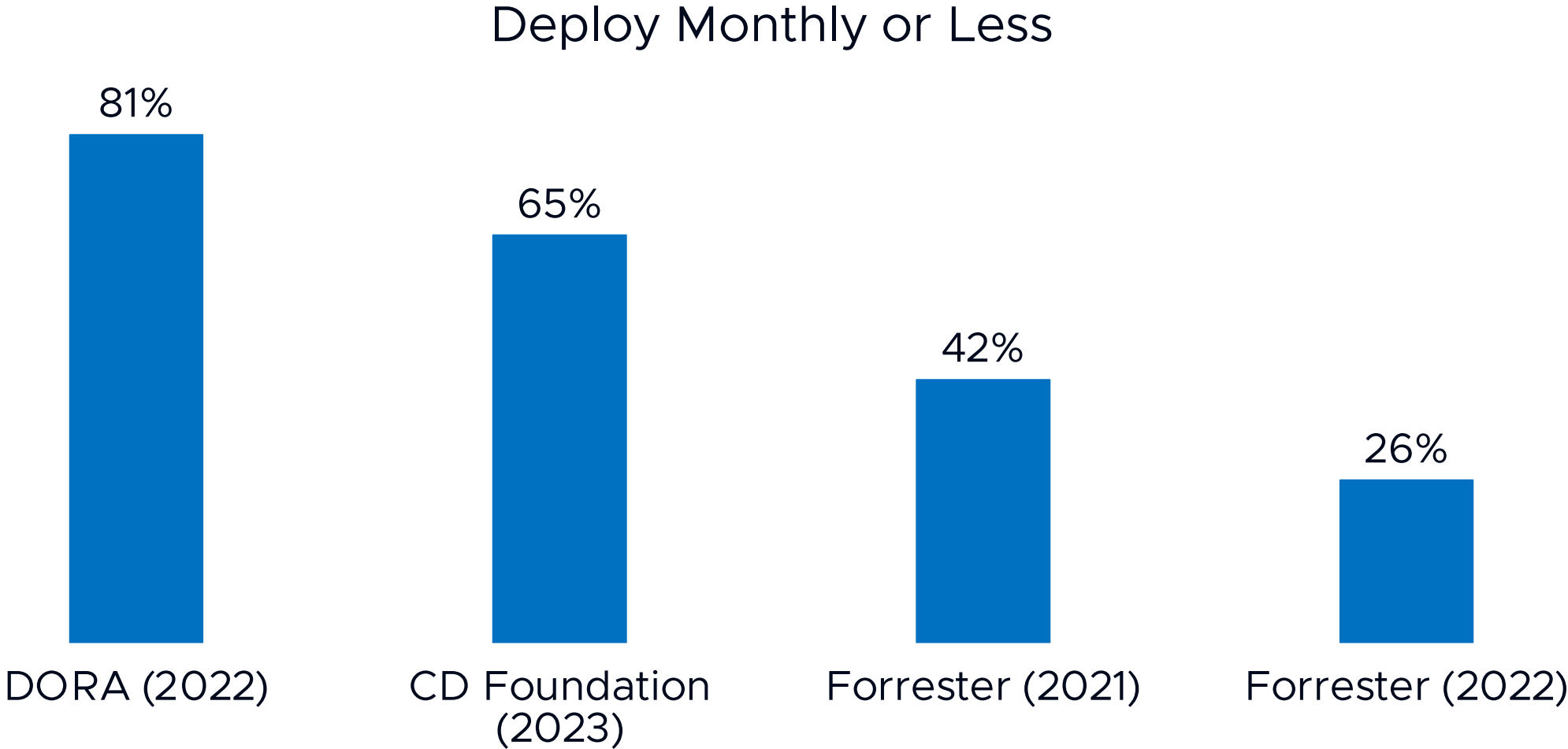
# Who is doing cloud native?

“By 2025, 15% of enterprise applications will run in a container environment, up from 5% in 2022, hampered by application backlog, technical debt, skills availability and cultural change.”

Gartner, CTOs' Guide to Containers and Kubernetes — Answering the Top 10 FAQs, Arun Chandrasekaran, Wataru Katsurashima, 31 May 2022

GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.

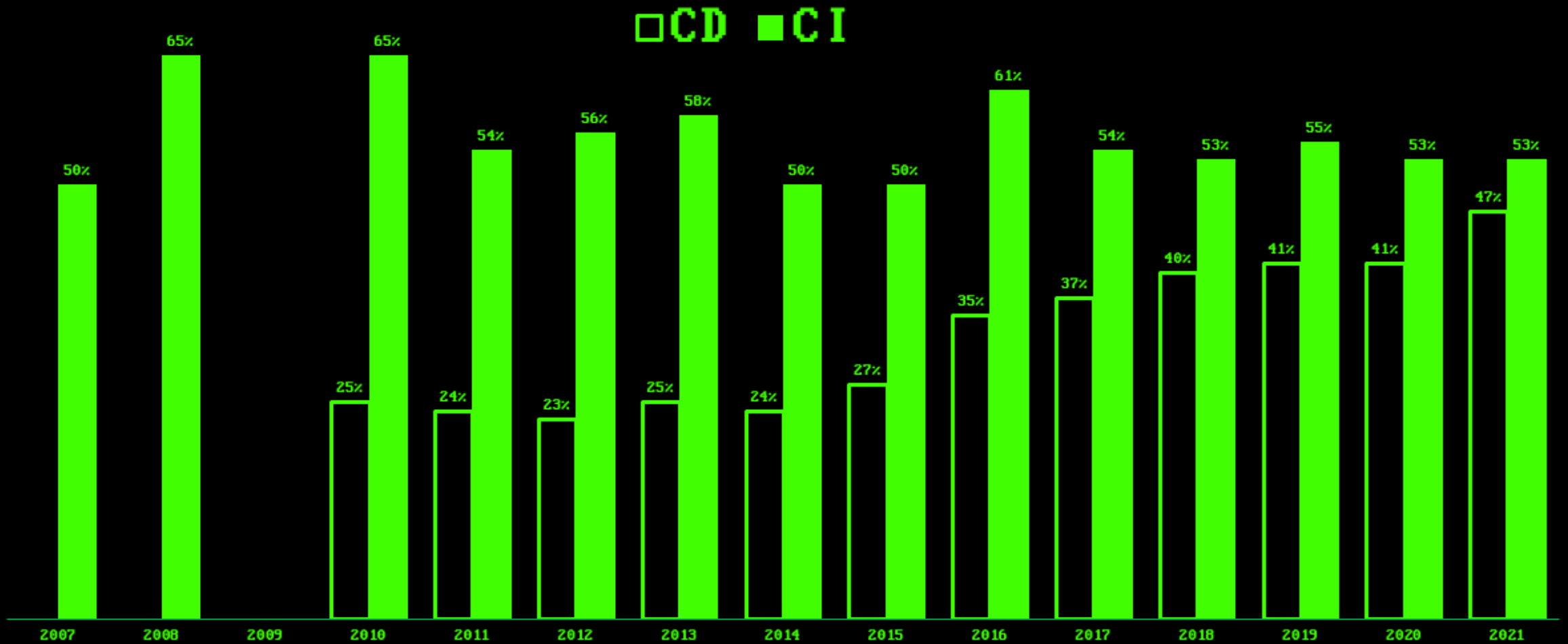
# Accounts of deployment rates vary wildly



Sources: see [“How's DevOps been going?”](#) Coté, June 2023 for citations and links to sources.



# CI and CD usage, 2007 to 2021



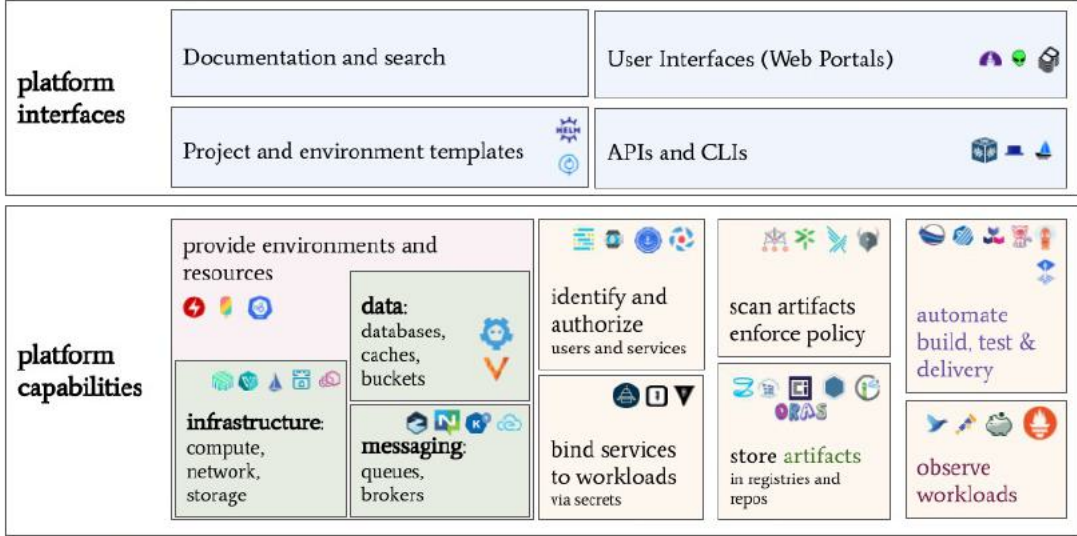
Source: State of Agile Surveys, VersionOne/CollabNet/digital.ai

#cote

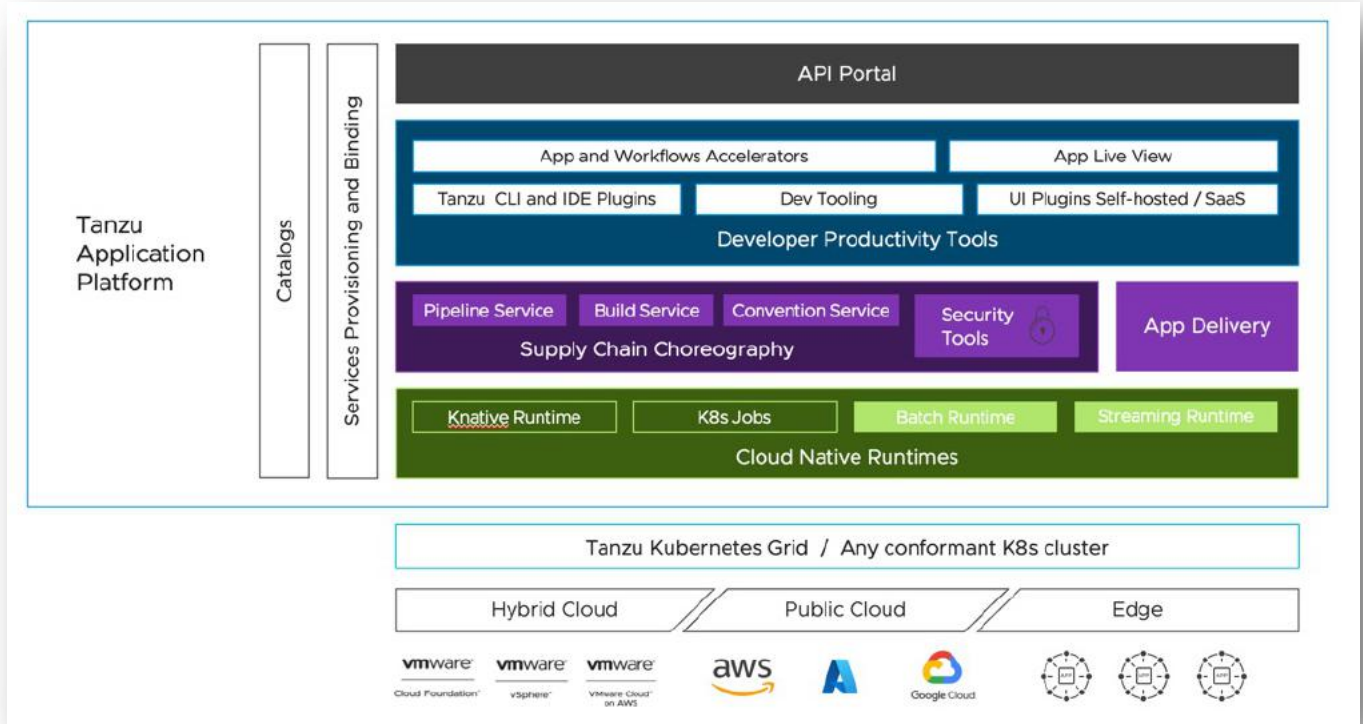
More: [“How's DevOps been going?”](#) Côté, June 2023.

# Platforms

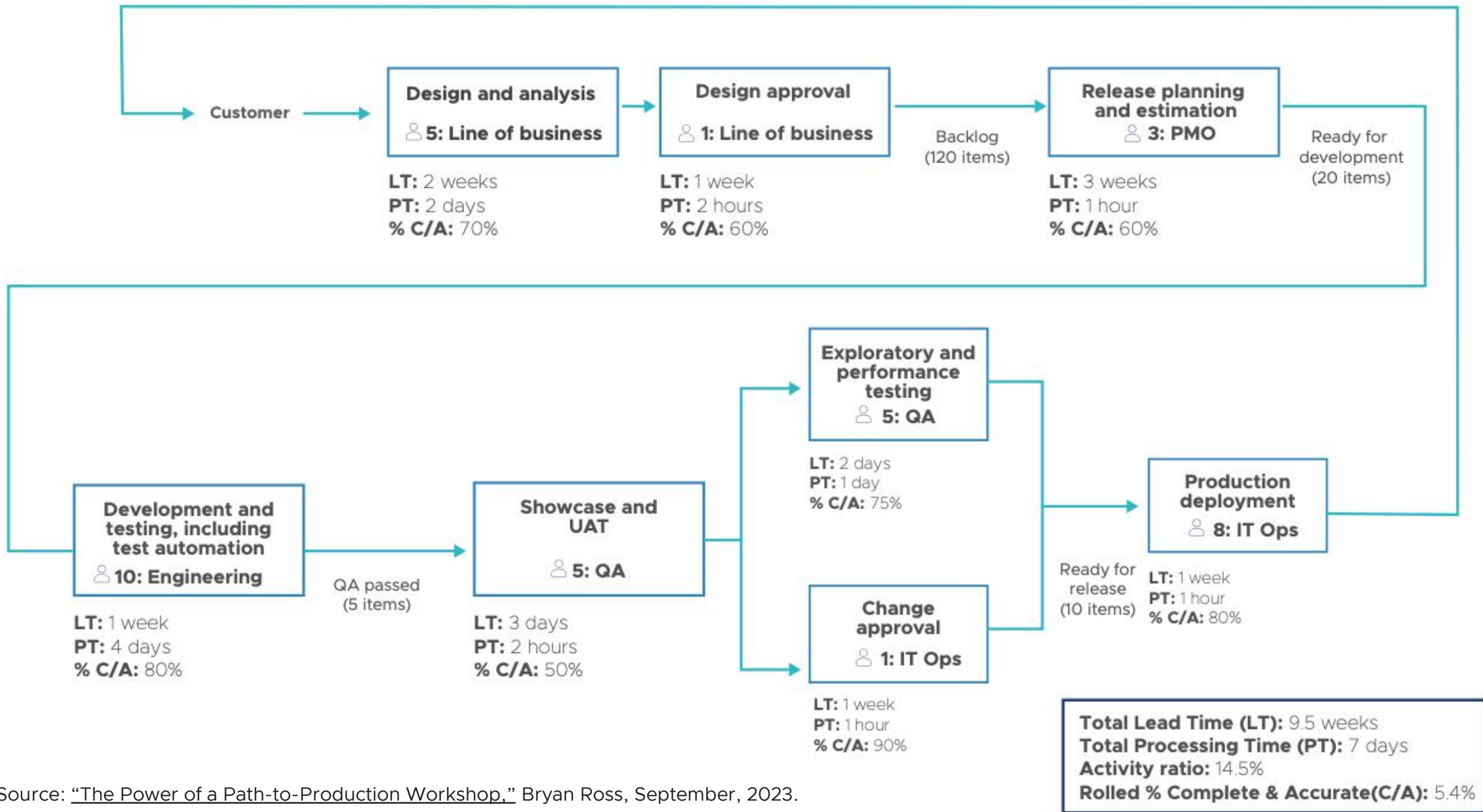
Product and application teams



Capability and service providers



Sources: [CNCF Platforms White Paper](#) and [Oct 2022 diagram draft](#); [VMware Tanzu Application Platform](#).



Source: "The Power of a Path-to-Production Workshop," Bryan Ross, September, 2023.

“We are building this platform not for us, we are building it for Mercedes-Benz developers.”

Thomas Müller, Mercedes-Benz



# Find the Developer Toil, Confusion, Blockers

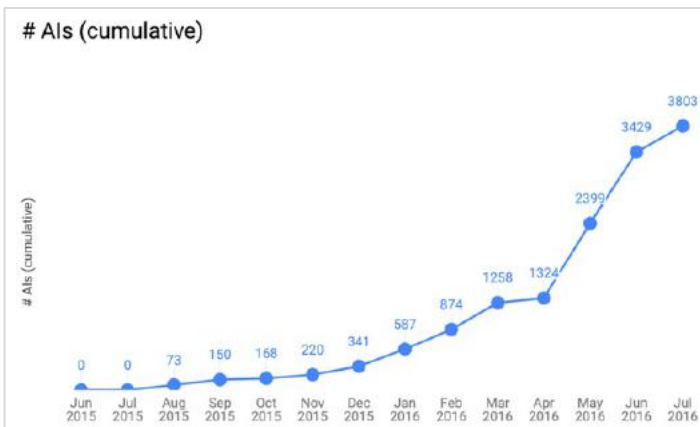
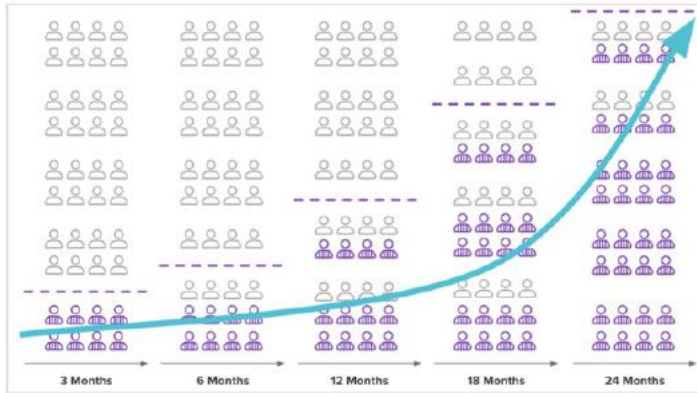
- What are we making?
- We have a strong vision for our product, and we're doing important work together every day to fulfill that vision.
- I have the context I need to confidently make changes while I'm working.
- I am proud of the work I have delivered so far for our product.
- I am learning things that I look forward to applying to future products.
- My workstation seems to disappear out from under me while I'm working.
- It's easy to get my workstation into the state I need to develop our product.
- What aspect of our workstation setup is painful?
- It's easy to run our software on my workstation while I'm developing it.
- I can boot our software up into the state I need with minimal effort.
- What aspect of running our software locally is painful? What could we do to make it less painful?
- It's easy to run our test suites and to author new ones.
- Tests are a stable, reliable, seamless part of my workflow.
- Test failures give me the feedback I need on the code I am writing.
- What aspect of production support is painful?
- We collaborate well with the teams whose software we integrate with.
- When necessary, it is within my power to request timely changes from other teams.
- I have the resources I need to test and code confidently against other teams' integration points.
- What aspect of integrating with other teams is painful?
- I'm rarely impacted by breaking changes from other tracks of work.
- We almost always catch broken tests and code before they're merged in.
- What aspect of committing changes is painful?
- Our release process (CI/CD) from source control to our story acceptance environment is fully automated.
- If the release process (CI/CD) fails, I'm confident something is truly wrong, and I know I'll be able to track down the problem.
- Our team releases new versions of our software as often as the business needs us to.
- We are meeting our service-level agreements with a minimum of unplanned work.
- When something is wrong in production, we reproduce and solve the problem in a lower environment.

Sources: "[Developer Toil: The Hidden Tech Debt](#)," Susie Forbath, Tyson McNulty, and Coté, August, 2022. See also Michael Galloway's [interview questions for platform product managers](#).

# Initial Phase – Create the Platform with Developers

1. Platform team: owner, operators, engineers, advocate
2. Pick one app for business & technical feasibility.
3. Developer toil audit.
4. Find path to production with end-to-end value stream analysis
5. Start with pre-integrated platform, customize as you...
6. Build a golden path with the developer team.
7. Optionally, build an IDP as part of the platform.
8. Do this for 3 months.

# Scaling Phase: Pairing & Seeding



1. Create platform marketing program.
2. Find two to five more apps.
3. Pair & seed from first dev & platform team to new teams.
4. "Shift Left" - build golden paths for governance, security, etc.
5. Add more infrastructure staff with pairing & seeding.
6. Do this for three months.
7. Repeat, growing number of apps as pairing & seeding allows.

Sources: "From 0 to 1000 Apps: The First Year of Cloud Foundry at The Home Depot," Anthony McCulley, The Home Depot, Aug 2016; "Cloud Native at The Home Depot, with Tony McCulley," Pivotal Conversations #45; USAF presentations and write-ups; "Driving Business Agility Without Large-Scale Transformation Programs," Venkatesh Arunachalam, Sep 2021; The Home Depot 2022[?]Q4 earnings call; *The Business Bottleneck*, Coté.





# Thanks!



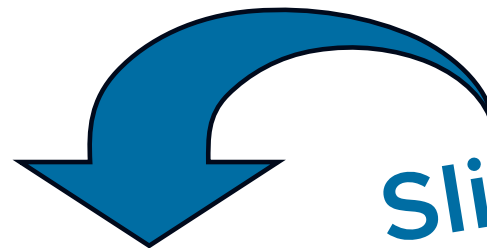
<https://newsletter.cote.io>



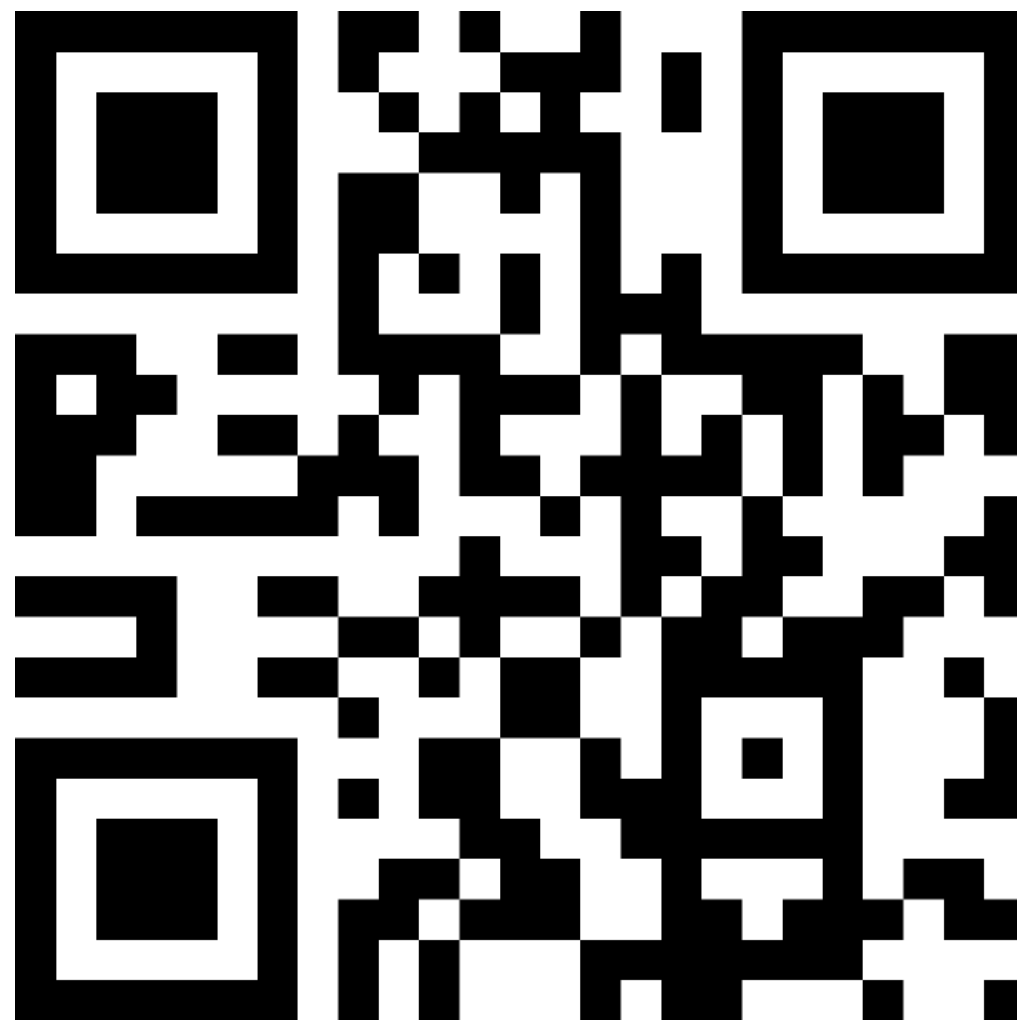
<https://cote.io/platform/>



[cotem@vmware.com](mailto:cotem@vmware.com)



slides & stuff





**ENTERPRISE  
DEVOPS**  
T E C H C O N