Perfecto
by Perforce

Best Practices to Overcome Continuous Testing in the Cloud Issues

SQGNE Software Quality Group of New England

Dedicated to Software Quality Professionals

Perfecto

Eran Kinsbruner

- Chief Evangelist, Product Manager, and Author at Perfecto by Perforce
- Blogger, Inventor, and Speaker
- 20+ years in software development & testing
- **Author of**:
  - *The Digital Quality Handbook*
  - *Continuous Testing for DevOps Professionals*
  - *Accelerating Software Quality*
- Twitter: @ek121268

THE DIGITAL QUALITY HANDBOOK
Eran Kinsbruner
A GUIDE FOR ACHIEVING CONTINUOUS QUALITY IN A DEVOPS REALITY
EDITION 1
CONTINUOUS TESTING FOR DEVOPS PROFESSIONALS
A PRACTICAL GUIDE FROM INDUSTRY EXPERTS
BY ERAN KINSBRUNER
ERAN KINSBRUNER
A Guide from key DevOps Industry Experts
Accelerating Software Quality
Machine Learning and Artificial Intelligence in the Age of DevOps
PERFORCE

Perfecto

Today’s Agenda

1 State of Test Automation

2 Top Challenges and Trends in Digital Apps

3 Best Practices in Continuous Testing

4 Live Demo and Q&A

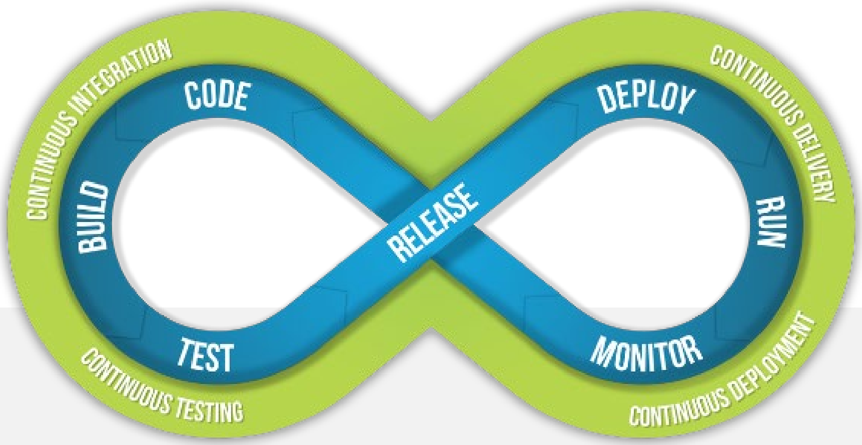
State of CT

tl;dr

State of Test Automation — 2021

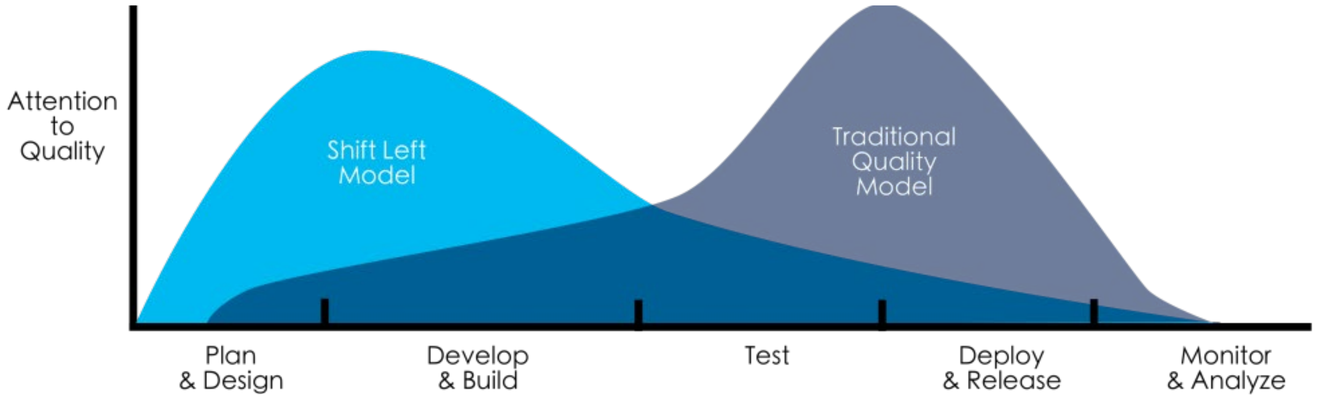
- **Teams focused on more automation, better coverage:** Despite link to faster releases, half of respondents automate fewer than 50% of tests. They're changing that in 2021.
- **Shifting left is shifting into high gear:** The movement is correlated with faster releases. Larger organizations lead the pack in adoption.
- **Top performers are addressing the automation skills gap.** With higher levels of automation within the release cycle, leading orgs are vetting BDD and scriptless solutions.
- **Leadership, practitioners not on same page:** C-Suite and VPs believe they are more automated than practitioners report. Alignment needed for prioritizing 2021 investments.
- **DevOps is often differentiated** by the following criteria:
  - Rapid release frequency.
  - Fully automated CI/CD pipelines.
  - Ability to promote/demote releases frequently.
- **Democratizing quality assurance across the entire organization** is key for DevOps success (shifting it **left and right**)
- **Growing practices helping shift-left** include:
  - Use of features flags (now supported by Android APKs as well).
  - A/B testing.
  - Cloud-based testing.

What Is Continuous Testing?



Continuous testing is the process of executing automated **high-value** tests as a part of the software delivery pipeline in order to obtain feedback on **business risks** associated with a software release upon every code change.

In Other Words – Modern Shift Left



Shift Left – Key Benefits

It's more **cost effective** to reveal defects earlier (Fail fast/Fix fast)

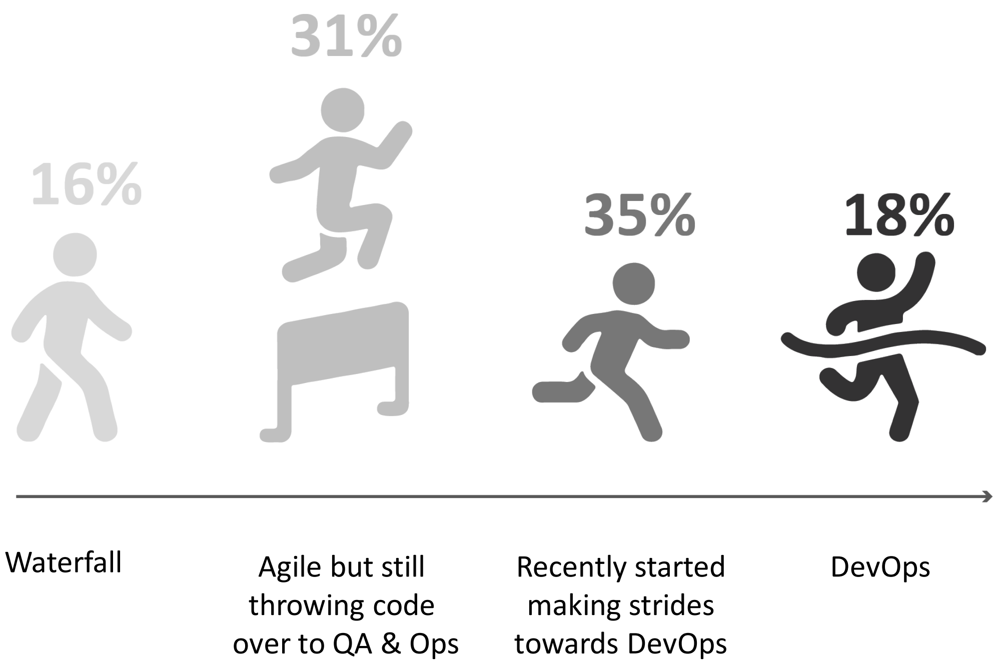
It supports the goals of **Agile/DevOps** (Releasing Faster)

Easier to **maintain code and quality**

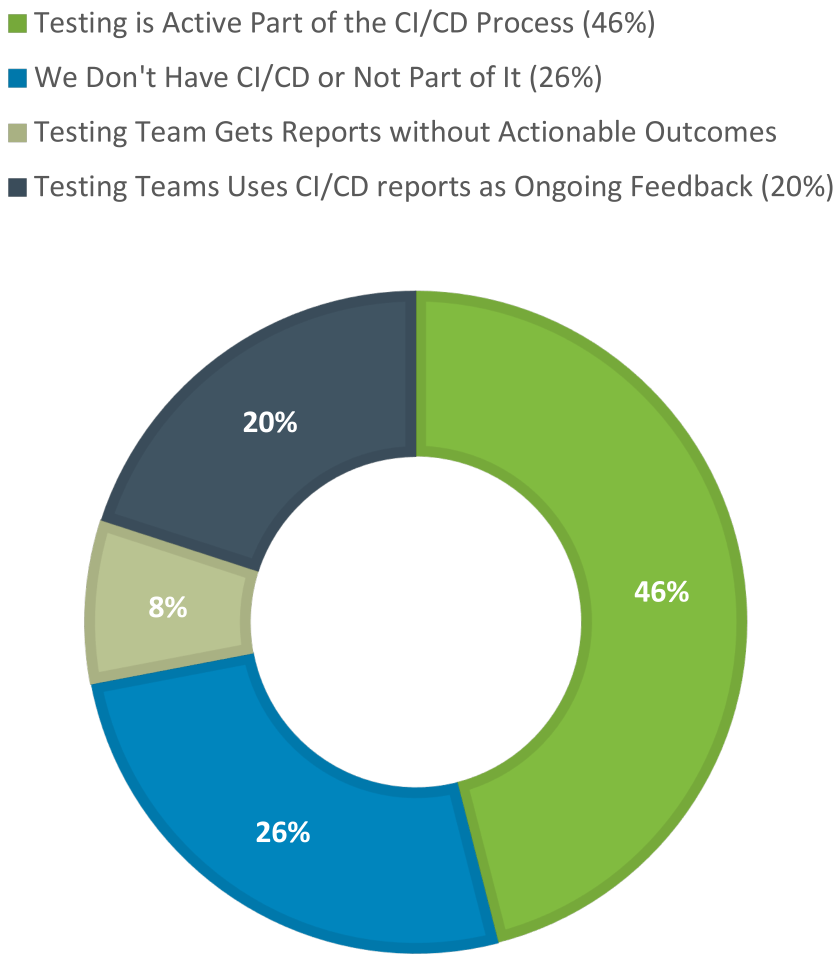
Cover more functional and Non-functional areas early (Security, accessibility, performance, API)

Testing As Part of the CI/CD Process

“WHERE ARE YOU IN THE DEVOPS TRANSFORMATION?”



TESTING ROLE WITHIN CI/CD



Top Testing Challenges and Trends for 2021

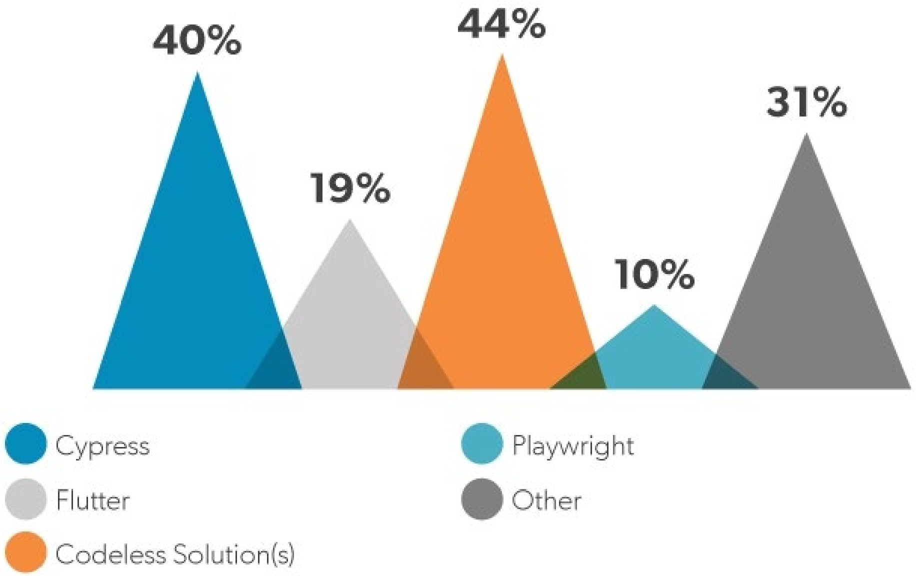
Top Software Testing Challenges

- Lacking resources for test creation
- Too many manual tests, inability to fit in schedule
- Achieving the right test automation coverage (what to automate, on which platforms)
- Automating advanced scenarios
- Testing not prioritized properly, done late in the cycle
- Deciphering why tests fail
- Matching tools to skillset
- Test flakiness (elements, etc.)
- Test environment setup & test maintenance
- Application requirements keep changing
- Justifying testing within the organization (proving value/ROI/budget)
- Understanding product requirements/user stories
- Working with Dev and Agile/TDD methods

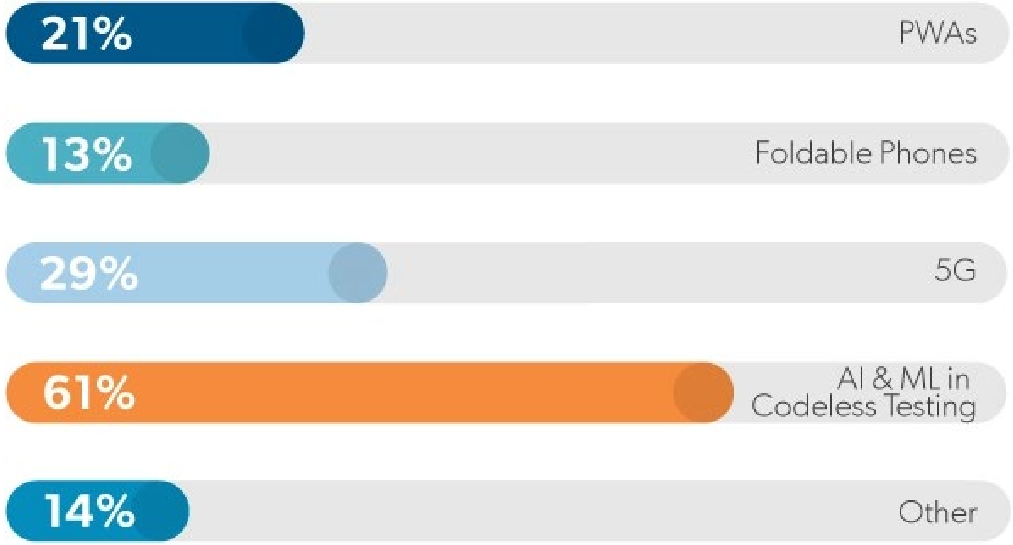
Perfecto

State of Test Automation — Looking Ahead

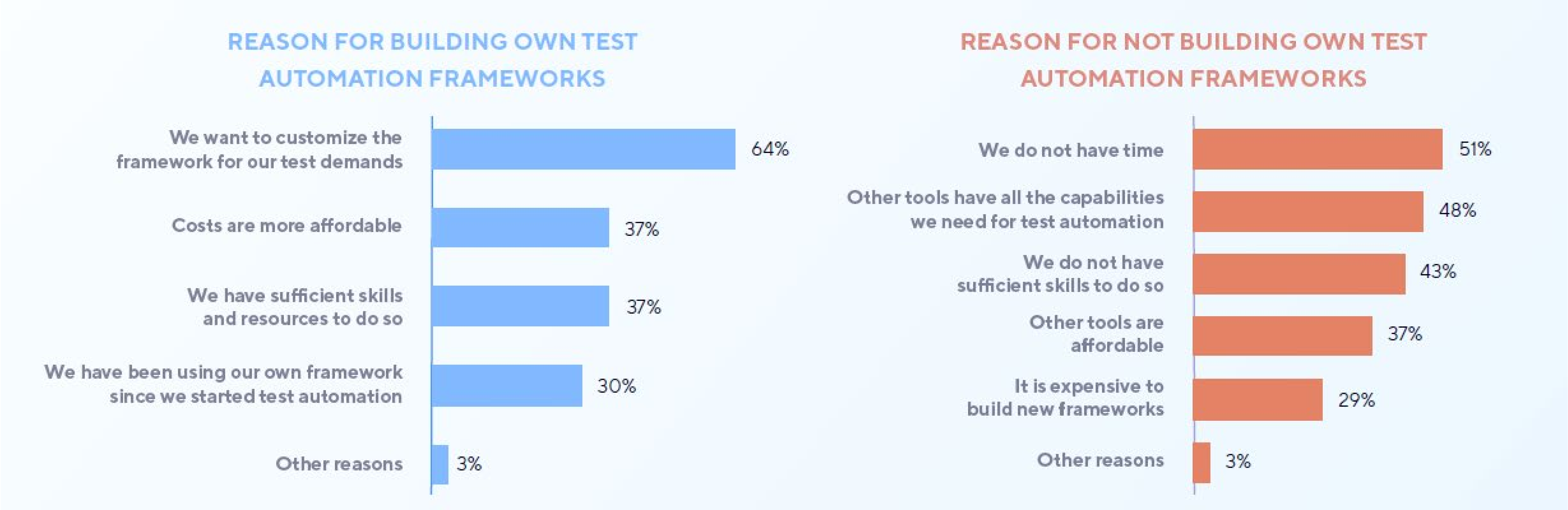
Which frameworks are top of mind to you in 2021? Select all that apply.



Which trends impacting testing are you looking to invest in this year? Select all that apply.

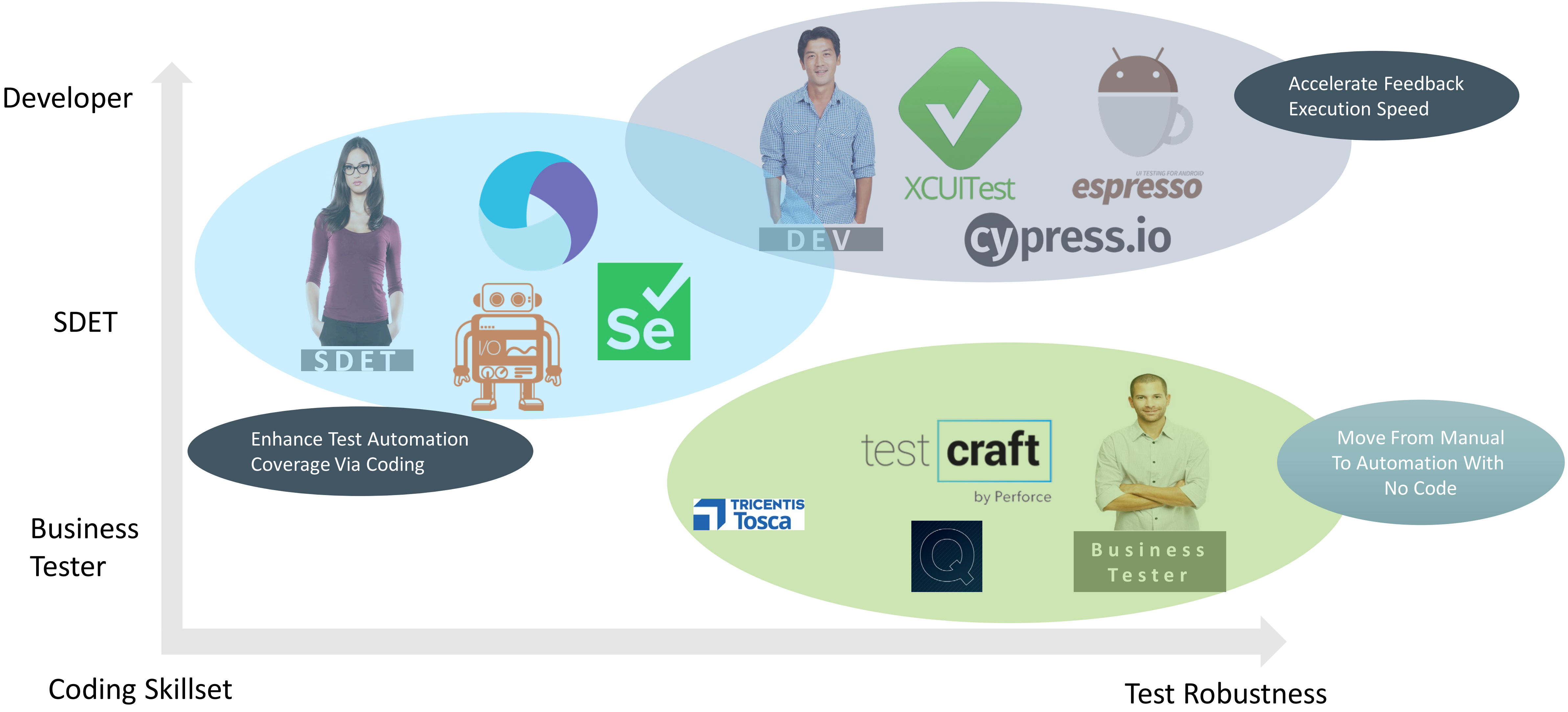


Why Are Teams Building/Not Building Custom Test Automation FWs?



Solving the Challenges

Matching Software Test Automation Technology to Org. Structure



Developer

SDET

Business Tester

Coding Skillset

Test Robustness

Accelerate Feedback Execution Speed

Enhance Test Automation Coverage Via Coding

Move From Manual To Automation With No Code

Test Automation Plan for an Airline App — Automating What's Right & of Value

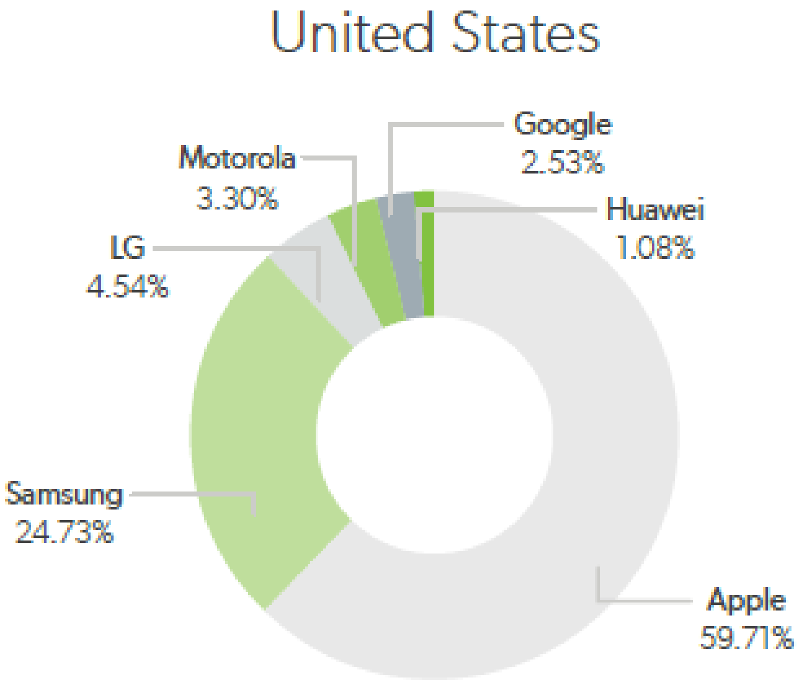
Distinguish various env. & real user dependencies throughout the journey:

- Locations
- Platforms (Web/Mobile)
- Networks
- Performance
- Accessibility
- Sensor Dependent Testing

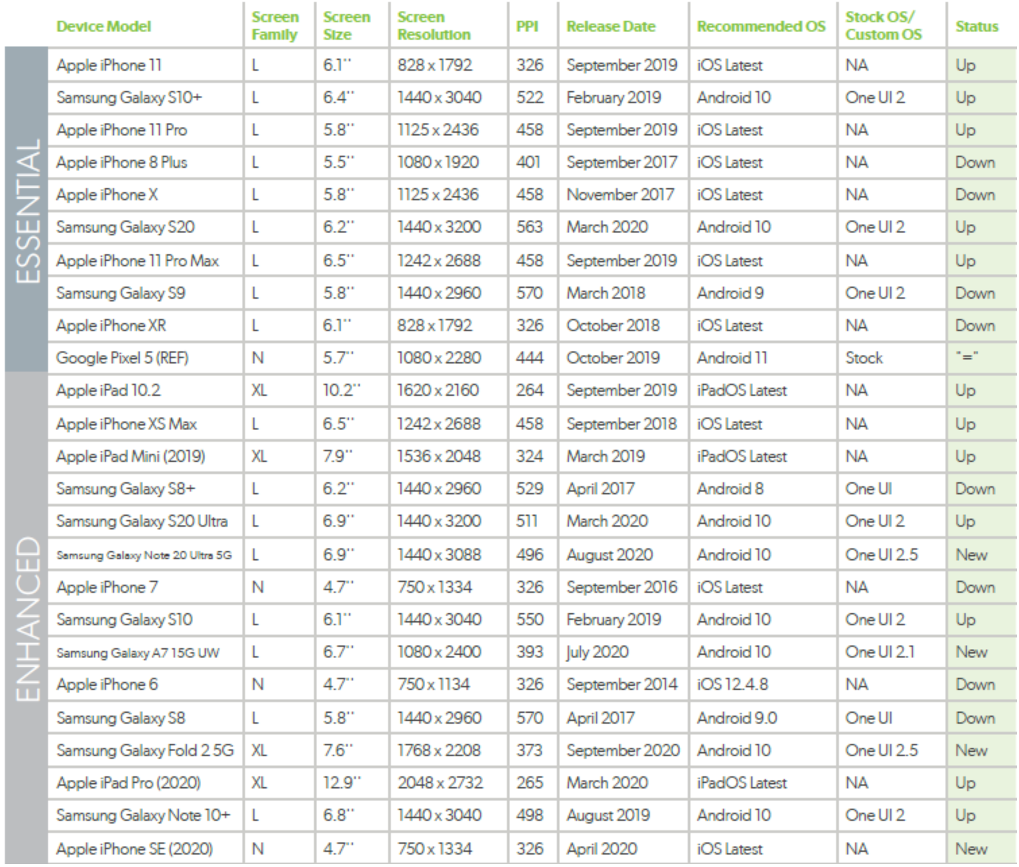
Feature	Critical / Freq.	Security	Data / Env. Dependent	Reuse
Login	☑	☑	☑	☑
Flight Status	☑	□	☑	□
Book a Flight	☑	☑	☑	☑
Contact Support	□	□	☑	☑
Check-in	☑	☑	☑	☑
Track my Luggage	☑	☑	☑	☑
Cancel Flight	☑	☑	☑	☑
My Bookings	☑	☑	☑	□
Special Offers	□	□	☑	☑
Accessibility Compliance	☑	☑	☑	☑
Seats Selection	☑	☑	☑	☑
Wallet Pass Upload	☑	☑	☑	☑
Airports Map View	□	□	☑	☑
Book a Car/Uber	□	☑	☑	☑
Track your plane for upcoming flight	□	□	☑	☑

Perfecto

Continuous Digital Coverage Tracking



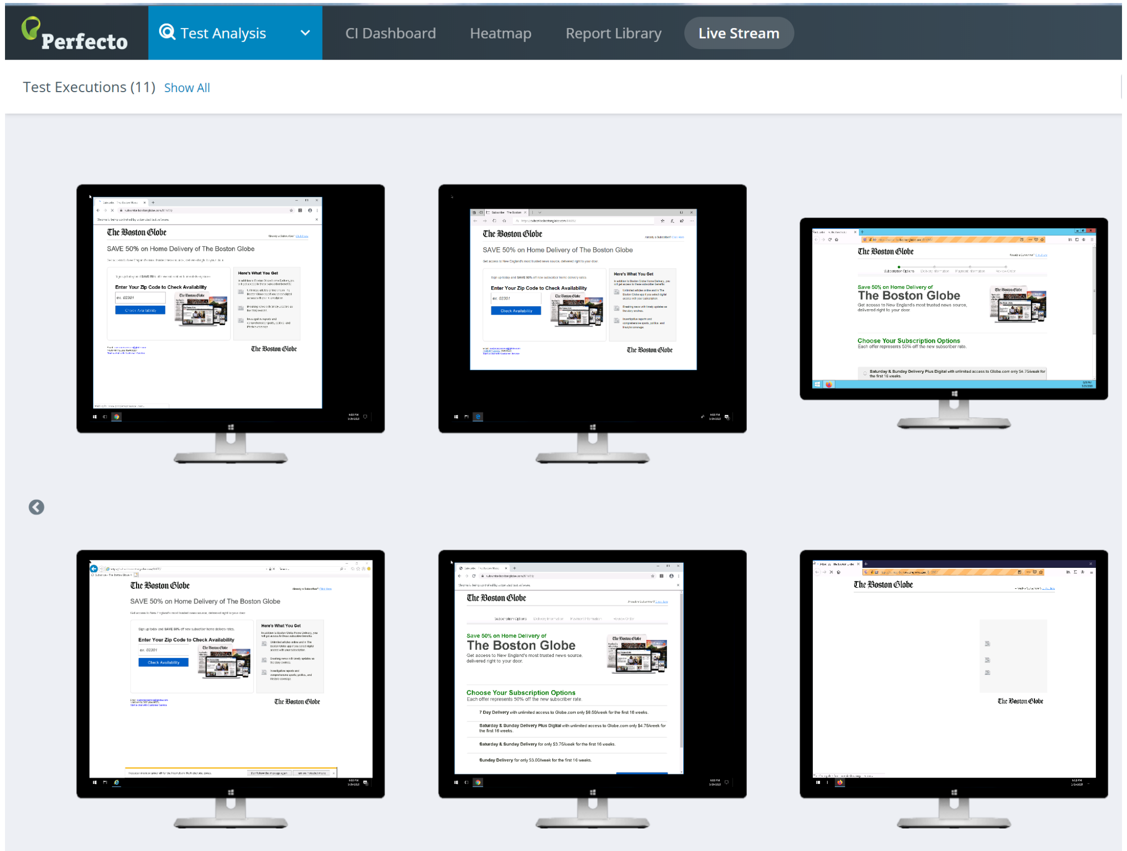
MOBILE DEVICE INDEX: US



	Device Model	Screen Family	Screen Size	Screen Resolution	PPI	Release Date	Recommended OS	Stock OS/ Custom OS	Status
ESSENTIAL	Apple iPhone 11	L	6.1"	828 x 1792	326	September 2019	iOS Latest	NA	Up
	Samsung Galaxy S10+	L	6.4"	1440 x 3040	522	February 2019	Android 10	One UI 2	Up
	Apple iPhone 11 Pro	L	5.8"	1125 x 2436	458	September 2019	iOS Latest	NA	Up
	Apple iPhone 8 Plus	L	5.5"	1080 x 1920	401	September 2017	iOS Latest	NA	Down
	Apple iPhone X	L	5.8"	1125 x 2436	458	November 2017	iOS Latest	NA	Down
	Samsung Galaxy S20	L	6.2"	1440 x 3200	563	March 2020	Android 10	One UI 2	Up
	Apple iPhone 11 Pro Max	L	6.5"	1242 x 2688	458	September 2019	iOS Latest	NA	Up
	Samsung Galaxy S9	L	5.8"	1440 x 2960	570	March 2018	Android 9	One UI 2	Down
	Apple iPhone XR	L	6.1"	828 x 1792	326	October 2018	iOS Latest	NA	Down
	Google Pixel 5 (REF)	N	5.7"	1080 x 2280	444	October 2019	Android 11	Stock	"="
ENHANCED	Apple iPad 10.2	XL	10.2"	1620 x 2160	264	September 2019	iPadOS Latest	NA	Up
	Apple iPhone XS Max	L	6.5"	1242 x 2688	458	September 2018	iOS Latest	NA	Up
	Apple iPad Mini (2019)	XL	7.9"	1536 x 2048	324	March 2019	iPadOS Latest	NA	Up
	Samsung Galaxy S8+	L	6.2"	1440 x 2960	529	April 2017	Android 8	One UI	Down
	Samsung Galaxy S20 Ultra	L	6.9"	1440 x 3200	511	March 2020	Android 10	One UI 2	Up
	Samsung Galaxy Note 20 Ultra 5G	L	6.9"	1440 x 3088	496	August 2020	Android 10	One UI 2.5	New
	Apple iPhone 7	N	4.7"	750 x 1334	326	September 2016	iOS Latest	NA	Down
	Samsung Galaxy S10	L	6.1"	1440 x 3040	550	February 2019	Android 10	One UI 2	Up
	Samsung Galaxy A7 15G UW	L	6.7"	1080 x 2400	393	July 2020	Android 10	One UI 2.1	New
	Apple iPhone 6	N	4.7"	750 x 1134	326	September 2014	iOS 12.4.8	NA	Down
	Samsung Galaxy S8	L	5.8"	1440 x 2960	570	April 2017	Android 9.0	One UI	Down
	Samsung Galaxy Fold 2 5G	XL	7.6"	1768 x 2208	373	September 2020	Android 10	One UI 2.5	New
	Apple iPad Pro (2020)	XL	12.9"	2048 x 2732	265	March 2020	iPadOS Latest	NA	Up
	Samsung Galaxy Note 10+	L	6.8"	1440 x 3040	498	August 2019	Android 10	One UI 2	Up
	Apple iPhone SE (2020)	N	4.7"	750 x 1334	326	April 2020	iOS Latest	NA	New

https://www.perfecto.io/resources/mobile-web-test-coverage-index

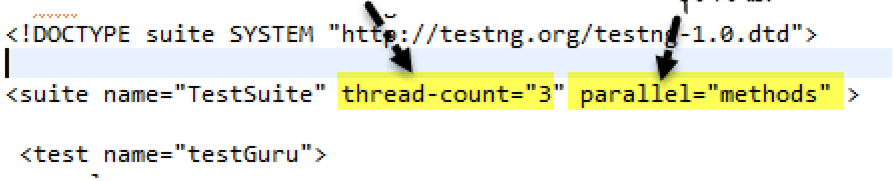
Parallel Test Automation



- Increased Coverage
- Faster Execution Speed
- Faster Feedback

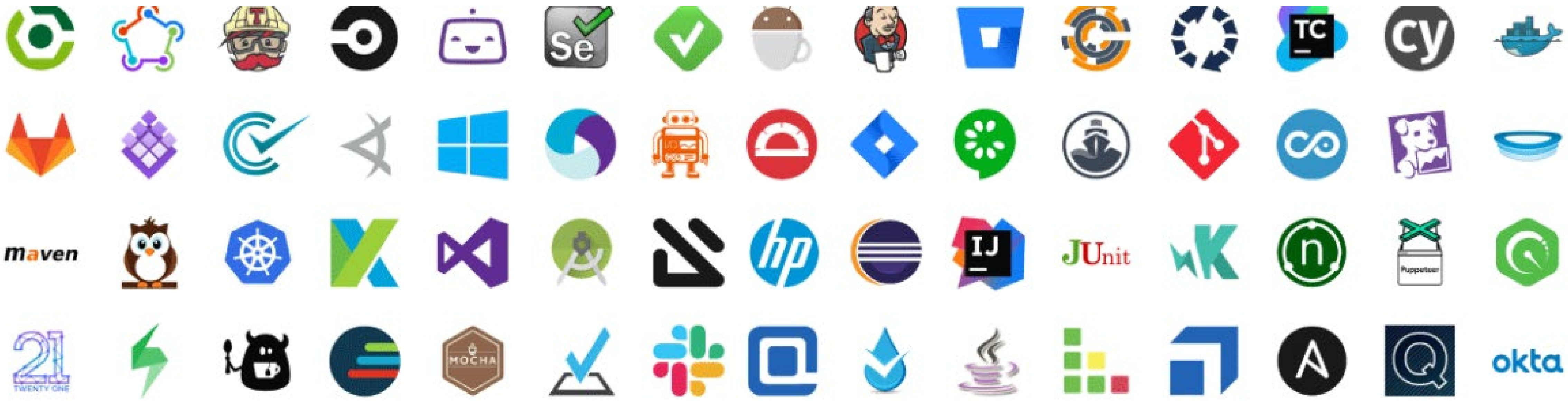
As the thread count is three, three WebDriver instances can run parallel in three different browsers

Test cases can run parallel



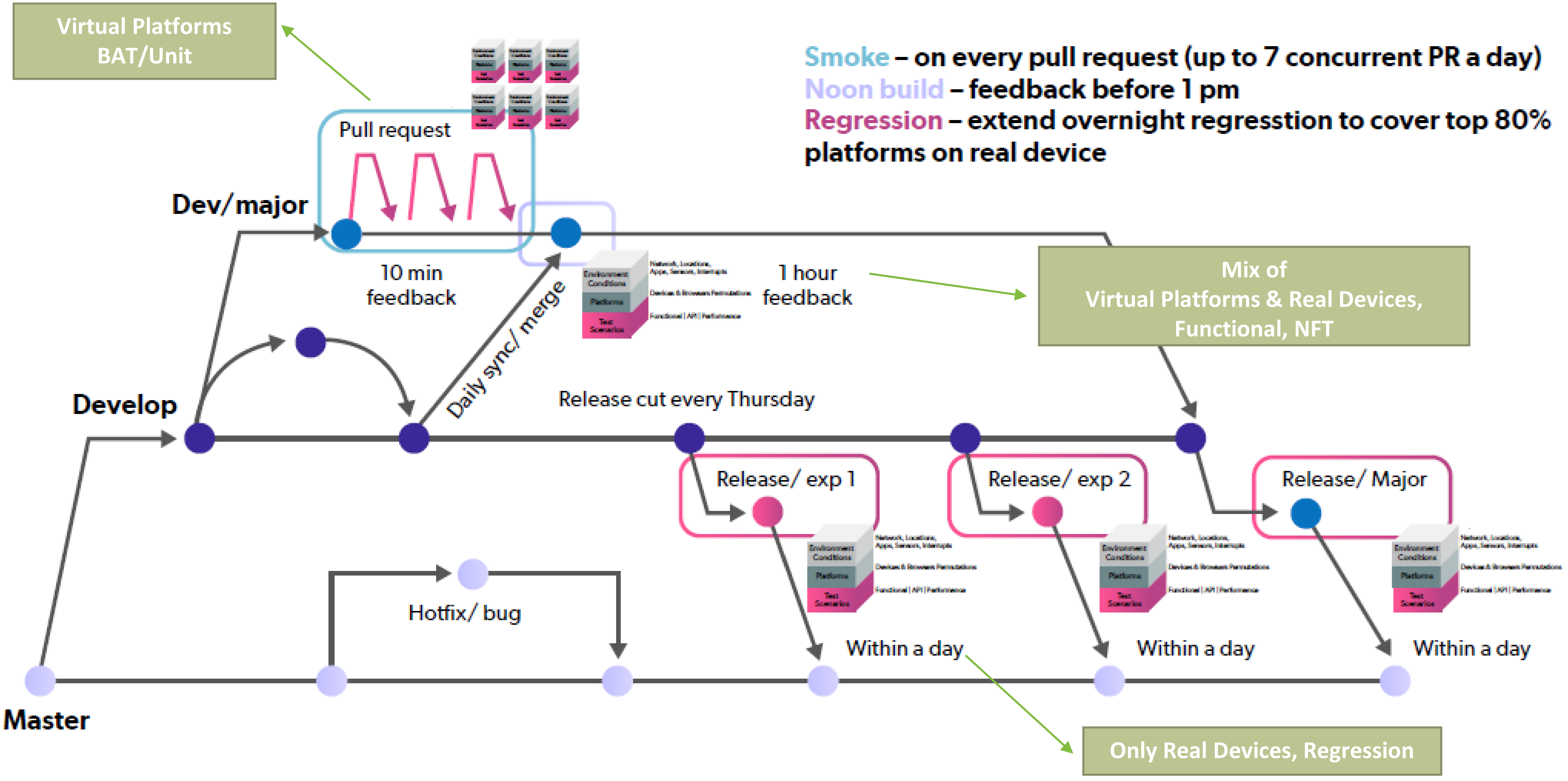
```
<!DOCTYPE suite SYSTEM "http://testng.org/testng-1.0.dtd">

<suite name="TestSuite" thread-count="3" parallel="methods" >

 <test name="testGuru">
```

Enabling Open & Integrated Reality



Keep A Solid Automated Pipeline

Smoke – on every pull request (up to 7 concurrent PR a day)
Noon build – feedback before 1 pm
Regression – extend overnight regresstion to cover top 80% platforms on real device


Virtual Platforms BAT/Unit
Pull request
Dev/ major
10 min feedback
Daily sync/ merge
1 hour feedback
Mix of Virtual Platforms & Real Devices, Functional, NFT
Develop
Release cut every Thursday
Release/ exp 1
Release/ exp 2
Release/ Major
Environment Conditions
Platforms
Test Scenarios
Network, Locations, Apps, Sensors, Interrupts
Devices & Browsers Permutations
Functional | API | Performance
Hotfix/ bug
Within a day
Within a day
Within a day
Master
Only Real Devices, Regression

PHASED COVERAGE APPROACH

Local Dev Team

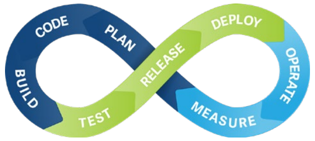
- Developers in teams should run smoke tests prior to commit
- 24/7 as needed basis to fix any bugs fresh in their mind.
- As a result less bugs are leaked and merged into main branch

Acceptance Test (Smoke)

- Support up to 9 teams and multiple branches
- 24/7 as needed basis
- Provide developer fast feedback to reduce the time spend on finding and fixing defects

Integration Tests (Sanity)

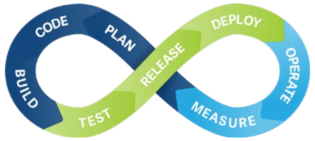
- 4 times a day
- Capture defects in fully integrated environment



Sanity

Full Regression

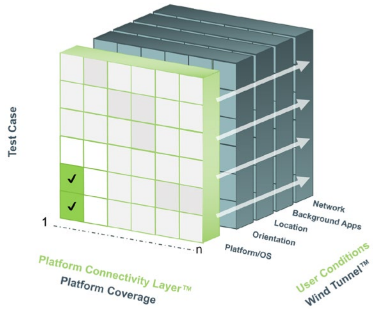
- Nightly
- Reduce defects leaking into production



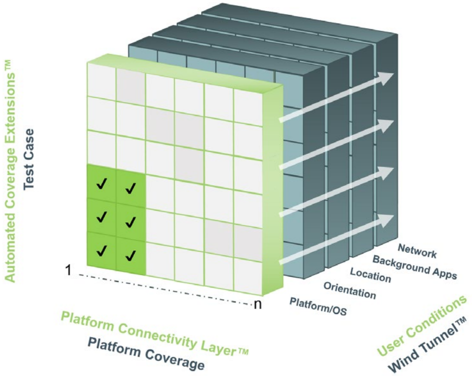
Regression

COVERAGE & FEEDBACK WINDOW

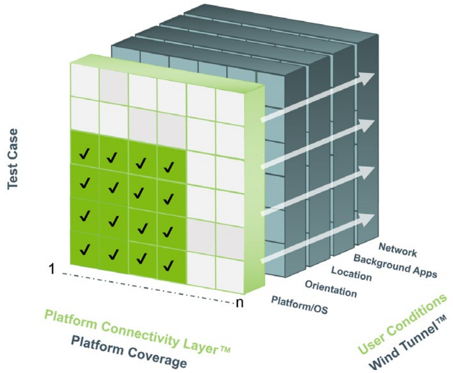
5-minute smoke tests
Validate on 1 device



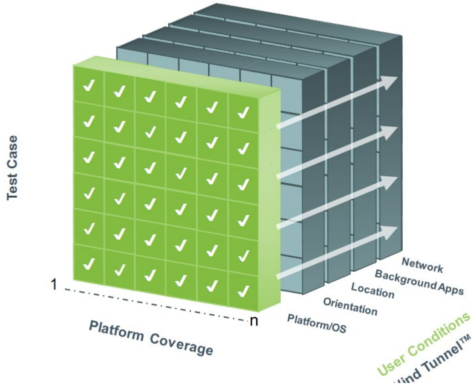
Every 5 minutes, smoke tests
1 android, 1 iOS



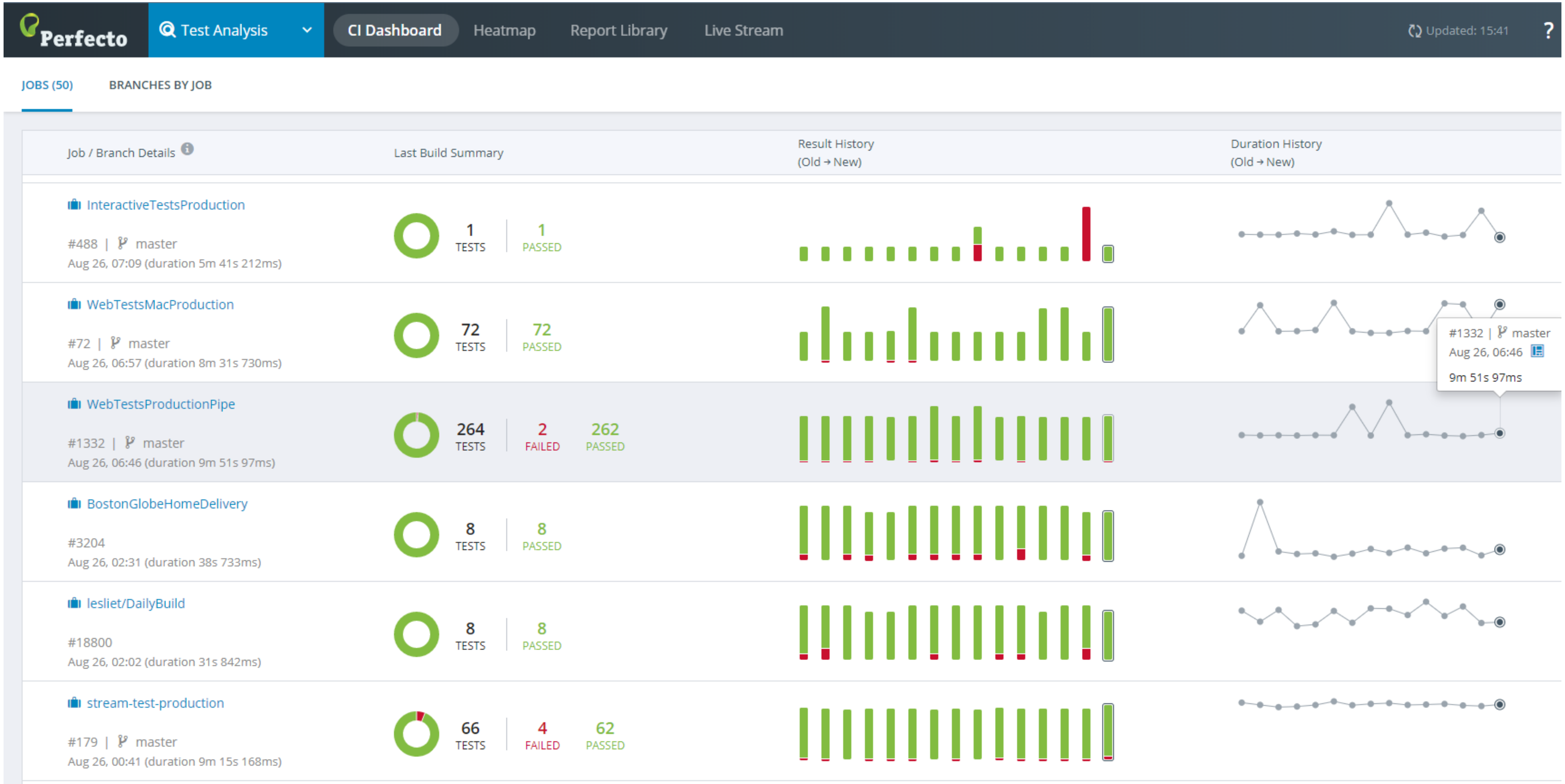
Every 2 hours, sanity suite
8 devices



Once a night (10 hours)
Top 32 devices



CI Pipeline Must be Clean & Green



Job / Branch Details	Last Build Summary	Result History (Old → New)	Duration History (Old → New)
InteractiveTestsProduction #488 | master Aug 26, 07:09 (duration 5m 41s 212ms)	1 TESTS 1 PASSED		
WebTestsMacProduction #72 | master Aug 26, 06:57 (duration 8m 31s 730ms)	72 TESTS 72 PASSED		
WebTestsProductionPipe #1332 | master Aug 26, 06:46 (duration 9m 51s 97ms)	264 TESTS 2 FAILED 262 PASSED		
BostonGlobeHomeDelivery #3204 Aug 26, 02:31 (duration 38s 733ms)	8 TESTS 8 PASSED		
lesliet/DailyBuild #18800 Aug 26, 02:02 (duration 31s 842ms)	8 TESTS 8 PASSED		
stream-test-production #179 | master Aug 26, 00:41 (duration 9m 15s 168ms)	66 TESTS 4 FAILED 62 PASSED		

#1332 | master
Aug 26, 06:46
9m 51s 97ms

Or Else: The Money Quickly Adds Up

Lost productivity from unstable build (in minutes)

== **Total Commits x Build Time X (1 – Stability %)**

<u>**EXAMPLE**</u>:

- 300 commits a week
- 30 minutes median build time
- 90% reliability on integration builds

300 commits * 30 Minutes * (1-0.90) = 900 Minutes a Week Lost!

Intrinsic Commit Build Cycle Cost

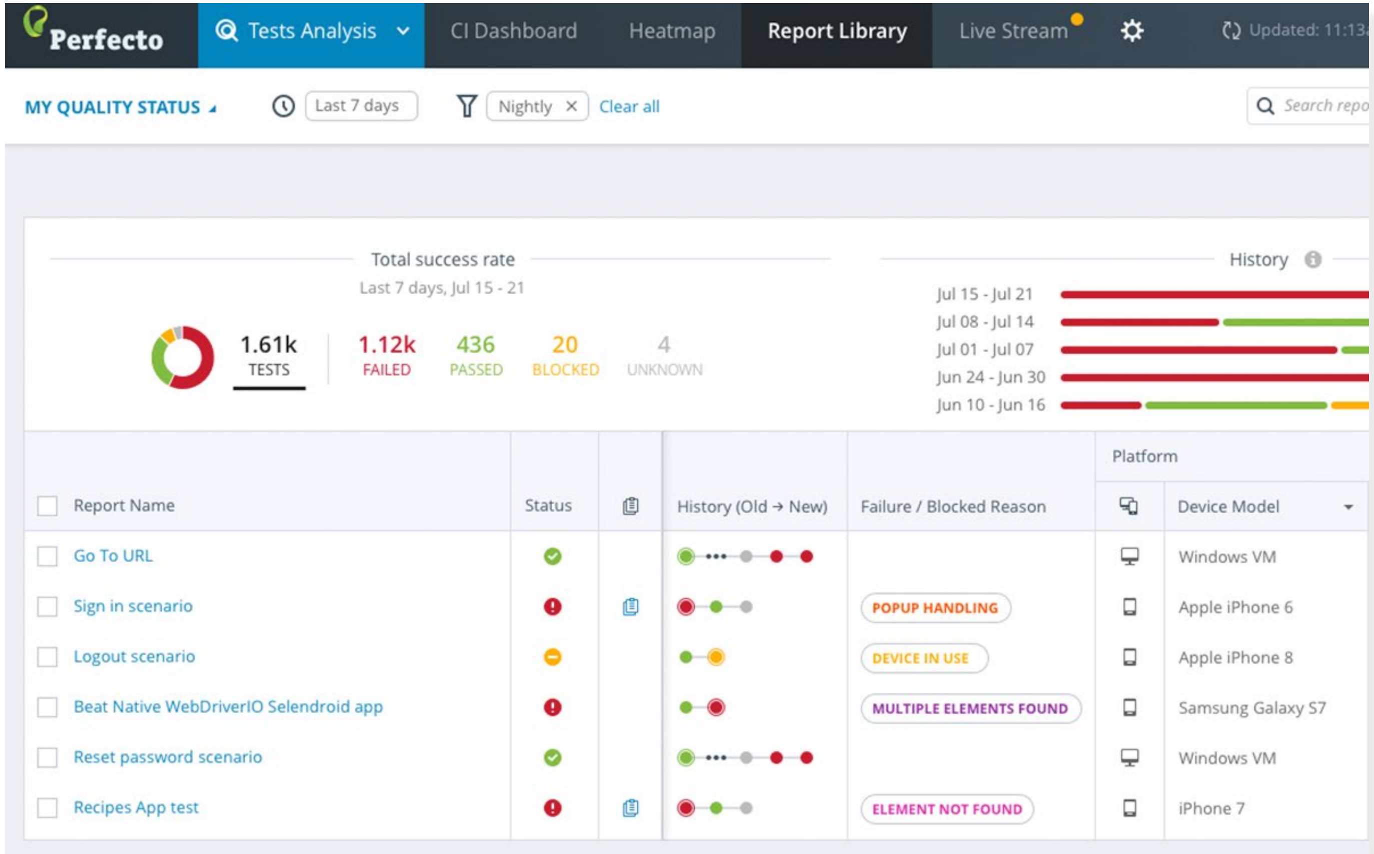
== **Build Run Time X (cost per machine hour + cost per developer hour)**

<u>**EXAMPLE:**</u>

- 25 Minutes median build time
- 5 minutes queuing time
- $1.00/hour machine cost
- $100/hour developer hour

0.5 hours * ($1.00 + $100) = $50.50 Cost/Build

What to Look at Within Test Failures?



Total success rate

Last 7 days, Jul 15 - 21

1.61k	1.12k	436	20	4
TESTS	FAILED	PASSED	BLOCKED	UNKNOWN

History

Jul 15 - Jul 21

Jul 08 - Jul 14

Jul 01 - Jul 07

Jun 24 - Jun 30

Jun 10 - Jun 16

					Platform	
□ Report Name	Status		History (Old → New)	Failure / Blocked Reason		Device Model
□ Go To URL	✓		● ••• ● ● ●			Windows VM
□ Sign in scenario	!		● ● ●	POPUP HANDLING		Apple iPhone 6
□ Logout scenario	−		● ●	DEVICE IN USE		Apple iPhone 8
□ Beat Native WebDriverIO Selendroid app	!		● ●	MULTIPLE ELEMENTS FOUND		Samsung Galaxy S7
□ Reset password scenario	✓		● ••• ● ● ●			Windows VM
□ Recipes App test	!		● ● ●	ELEMENT NOT FOUND		iPhone 7

Ask yourself:

1. Is there a long history of failures:
   - A. Test **flakiness** ?
     - A. If Yes: **Exclude** and **debug**
2. Are we **following** script development **practices** to avoid:
   - A. Popup/security issues
   - B. Element not found
   - C. Platform-specific issues
3. Do the tests bring any new **value**?
4. **Platform/CI health**?
5. Treating test code as production code? Versioning? Tagging?

Perfecto

Continuously Measure & Improve!

Test Automation Maturity

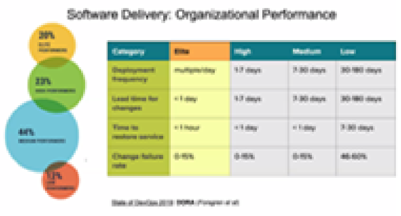
- Testing Types **Automated**: **% vs. Entire Suite**
  - Unit
  - BAT
  - Integration
  - Functional/UI
  - API
  - Regression (60% Automated, 60/100)
  - Performance
  - Accessibility
  - Security
- Test Automation **Creation**:
  - Time to create new test
  - Average test scenario execution time
- Test Automation **Reliability**
  - % of tests blocked
  - % of tests finding defects
  - % of tests flaky
  - % of tests constantly failing

Continuous Testing Visibility (Are we in control?)

- Platform-specific issues dashboards
- Recurring failures/defects
- CI dashboards
- Test suite heatmaps

CI/CD Pipeline Health/Dev Productivity

- % of tests per type running from CI, per code commit
- Average number of broken builds per software iteration (Build Pain Index)
- Average test failures per CI build within software iteration
- Average time for BAT
- Number of escaped defects found outside of CI (defect leakage)
- Productivity of developers — commits per day
- Release planning accuracy



Software Delivery: Organizational Performance

20%
Elite performers
23%
High performers
44%
Medium performers
12%
Low performers

Category	Elite	High	Medium	Low
Deployment frequency	multiple/day	1-7 days	7-30 days	30-180 days
Lead time for changes	< 1 day	1-7 days	7-30 days	30-180 days
Time to restore service	< 1 hour	< 1 day	< 1 day	7-30 days
Change failure rate	0-15%	0-15%	0-15%	46-60%

State of DevOps 2019 DORA (Forsgren et al)

Tools Match to DevOps Teams

- % of manual tests (tests that cannot be automated)
- % of false positive/negatives per build
- # of defects rejected by dev per software release
  - Cannot recreate
  - Not a defect
- Cross-team autonomy

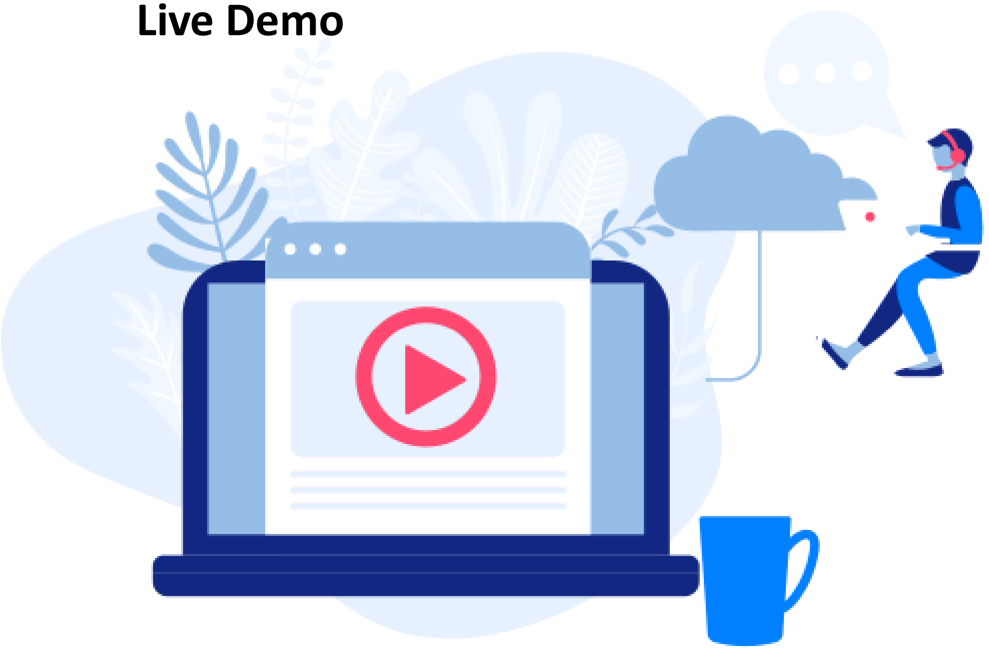
Test Coverage

- Platform coverage → # of escaped defects
  - Up to date
  - Sufficient
- Number of escaped defects found outside of CI

Product Quality/Other

- Mean time to resolution of production defects
- Manual activities within software iteration
- Customer feedback (App Store ratings/App Annie), # of bugs found by customers (monthly)
- Bug fixes deployed to production per week/per software iteration

Live Demo



Thank You!

Visit perfecto.io or follow us @perfectomobile