Testing in Production

Good, Bad or Ugly idea

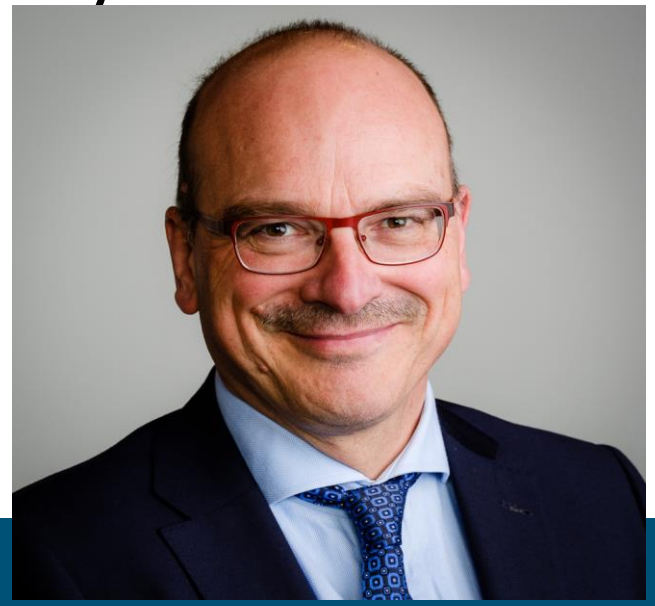
Jos van Rooyen

identify
THE QUALITY DRIVER

Agenda

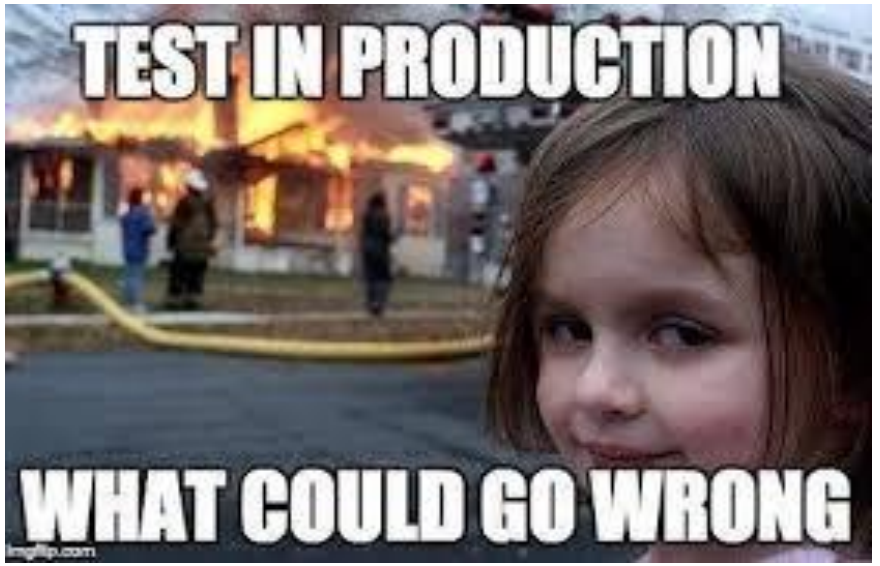
- Introduction
- Reasons to test in production
- Definition
- Forms of testing in production
- Areas of application
- Risks
- QA and production testing
- Pros and cons
- Examples
- Future work

- Employed at Identify as partner / principal consultant
- 30 years in software testing & quality management
- Co-author TestGrip, TestFrame, Project de Baas, Quality Supervision, Textbook; “Aan de slag met software testen”, Cleantxt, Test Automation Architecture
- Test expert online magazine Computable
- Publication areas; Test process Improvement, BI-testing, Test automation, Test Education, Risk Based Testing, Quality Supervision
- Visiting lecturer Universities of Applied Science
- Member advisory board Hogeschool Utrecht
- Board member of Dutch Testing Society



Introduction

- Up until a few years ago testing in production was almost not allowed. The test methods used didn’t stimulate it
- The idea was, it was too risky. The impact could not be overseen
- Testing doesn’t stop anymore by go live
- Bitter necessity
- It is not a taboo anymore



Challenges in software testing

- Systems are getting much more complex
- Too expensive to create a test environment
- Impossible to validate systems in development environment only
- Specific data situations are not available
- Historic data is not always available. Testing in the past is not possible

Reasons to test in production

- Time to market
- Complexity of the software is increasing. Much more components, interfaces are involved
- Too expensive to create a special test environment. Not really a quality issue but a budget issue
- Preparation for a new product or service
- Determine customer behaviour
- Not able to create the necessary circumstances
- Required test data not available in the test environment

Definition

Execution of test scenarios for specific situations in production because conditons could not be fullfilled during development proces(s). Based on a risk assessment and limited to certain branches or applications.

Types of testing in production

Feedback in production

- Monitoring and alerting
- Analytics events
- Logging
- Customer feedback

Test practices in production

- A/B testing
- Beta testing
- Monitoring as testing

Exposure control

- Canary release
- Staged rollout
- Dogfooding
- Dark launching

Areas of application

- Not always applicable due to several reasons:
  - Risk factor
  - Mission critical environment
- Applicable to:
  - Testing of workspace hardware
  - Automotive situations
  - Data input situations
  - Validation of customer behaviour

Risks

Risks of applying testing in production:

- Immature quality of the system
- Uncertainty of the chain. Not everything is foreseen
- Roll back is not possible
- Accidents or people get harmed
- No fallback scenario available

QA and production testing

- Execute risk assessment with all stakeholders
- Determine the situations for production testing
- Start testing during development
- Start with a small group (pre pilot, pilot, department, etc.)
- Scale up when possible but be aware of the risks
- Differences with a regular test process:
  - Real life situations are used
  - Negative testing is not possible. Only happy flow can be applied
  - Risk analysis upfront with all stakeholders and participants
  - Suppliers must be available 24/7
  - Focus on functional testing instead of also Non Functional Testing
  - Production testing is an extension to the test process

Pros and cons

- Pros:
  - Extra quality measures before implementation
  - Insight of impact several choices
  - Secured and staged rollout
  - Fast feedback loop
- Cons:
  - Too much risk
  - High costs
  - Context of the situation
  - Think about the environmental issues

Examples

Migration Citrix to O365 workspace	
No. Of workspaces	800
No. Of applications	220 (SAP, cloud etc.)
Challenges	Lack of knowledge of the applications Keyusers are brand new Sources not always available Usage not always clearly
Solution	Test in production by hand of a new device Citrix enviroment still in place as a back up
Results	250 bugs found and solved Succesful migrations all 18 locations

Examples

Innovation centre	
No. Of workspaces	
No. Of applications	
Challenges	
Solution	
Results	

Examples

Upgrade anti virus software	
No. Of workspaces	800
No. Of applications	200
Challenges	Too expensive to organize a test project IT team classify the risk as very low No testscripts available
Solution	Test by hand of 2 groups: It team test the software during a week just during normal business Next step is during a month with 60 users just doing there regular work When all these steps are successful the whole organization will work with the new antivirus software
Results	?

Applying the SAFe model

Case Study	Industry	type project	Challenge	Related costs to organize	Benefits
Tax office	Overheid		Tax office sought to improve its speed and predictability in bringing new technology to the organization and citizens	not available	**Technical Debt Down 80 Percent** Tax office continues to run two large ARTs (125+), with four Value Streams (one in Tax Allowances and three in Interaction Services). In fact, Belastingdienst follows a hybrid way of working where every department can choose either SAFe or a more 'traditional' project management-oriented way of working, depending on what fits best. The organization has driven notable results across the two ARTs and within a few smaller ARTs: **More frequent releases** – Major releases come out 3X more often, from 4 to 12 in a year. **Improved software quality/technical debt** – DTCA improved quality by reducing the number of 'problems' by 80 percent, and security issues by 87 percent (Interaction Services). **Less management overhead** – The number of people with the word 'manager' in their titles dropped in half. These individuals moved into other roles. **Increased engagement** – People are more engaged, connected with each other, and willing to help others.
Air France-KLM	Luchtvaart		Air France - KLM sought to scale Agile practices companywide to improve time to market and efficiency, but must contend with specific contexts and regulations in the different businesses of the airlines.	not available	**Results: 20% More Effective Delivery** Since deploying SAFe, Air France – KLM notes greater collaboration between business domains and Transversal Tracks. Within three months, their efforts began paying off in business results in the Cargo group: **Time-to-market** – Each ART team delivers on its promises every three weeks. Since moving to SAFe, the company released 17 times in the live environment in seven months compared to every six months previously. Quality – Of the 17 releases, the company had to delay just one due to a major incident **Productivity** – SAFe teams deliver, on average, more than 20% more effectively than waterfall teams **Adaptability** – With a PI cycle of 12 weeks, Air France – KLM has been able to pivot its vision three times in the past year, allowing the company to tap into new business opportunities much more quickly and easily **Market share** – The company gained 20% market share in the small and medium logistics market alone with this flexibility **Predictability** – The velocity of ARTs builds in more predictability and enables teams to take ownership and show greater craftsmanship. Team stability is also an important success factor in results **Business value** – On one offering, the company exceeded expectation by 25% Employee satisfaction – PI Planning results in better transparency and autonomy for the teams. Seeing the vision in the Cargo group encourages team members to contribute to the business value and increases their work satisfaction, as well as collaboration between business and IT **Customer satisfaction** – Air France – KLM is more intimate with its clients. All Product Owners from the business side have a greater understanding of the demand. Going live with small changes and new functionality every three weeks gives them a faster feedback loop and more rapid pivoting, enabling groups to deliver greater value in its IT solutions
Deutsche Bahn	Transport		After privatizing the company, Deutsche Bahn faced new market forces, along with increasing competition from new transportation players.		**Lead Time Down by 2/3** DB now runs 15 ARTs with four Value Streams, with seven of those at DB Netze and one in DB Cargo, five in DB Vertrieb, and at least two in other business units. Because of its efforts, DB now has MVP-based delivery, and manages dependencies within and between trains to avoid bottlenecks. In total, DB has trained more than 1,000 people in the past three years and has seen its dedication pay off: **Faster time-to-market** – Lead time dropped from 12 months to 3-4 months **More test automation** – Coverage of test automation improved from 30% or less to 80% **Better engagement** – 90% greater collaboration among teams and the resulting better outcomes have raised employees' satisfaction levels. Even skeptics now say, "Don't change it. **Clearer fiscal visibility** – They improved forecasting for financial requirements **Greater transparency** – DB has a crisper view on the portfolio roadmap, work in progress, and financial resources allocation
Cerno	Information Technology, Software		Deliver custom solutions faster and with higher quality for clients.	not available	**Delivery Cycle Time Down 58 Percent** Today, Cerno runs two ARTs with 80 people. These high-confidence teams agree on, and begin working on, requirements faster. They communicate and collaborate more tightly than before they introduced SAFe and are continuously improving. When the ART completed work with one client, they simply switched the train to support another logistics client with a similar solution—effectively a plug-and-play release train. The company then added a second ART to deliver value to another client. Each train continues to serve a single client. To date, Cerno has made remarkable progress: **Cycle time decreased** - Delivery cycle time dropped from 3½ weeks to two weeks, or 58 percent **Offline time decreased** - The average offline time for a new production environment release decreased from 3½ hours to half an hour **Failures went down** - The rate of release failure went down from 0.6 times on average per release to 0 **Automation increased** - The interface automation level increased from zero to 70 percent **Defects decreased** - Reported defects decreased from 13 times per release to five
Murex	Financial		With its MX.3 platform in use across the globe, Murex sought to maintain and build upon its market-leading position while continuing to respond rapidly to support the changing needs of clients and global regulatory demands.	not available	**Impressive Productivity Gains** As of today, Murex has trained more than 1,000 people in SAFe, or half the company, with teams distributed across its three development centers in Paris, Dublin, and Beirut. Its efforts have driven measurable progress across numerous benchmarks: **10X faster production-like testing** – Client Delivery teams can now simulate 10 weeks of real production activity in a single weekend **Complete testing in just one hour, instead of days** – The full client delivery testing cycle, including environment provisioning, functional tests, and upstream/downstream interface validation dropped from five days to just one hour, making it possible to run this full suite to customize each new customer configuration **85% reduction in user story cycle time** – Internal user story cycle for MX.3 platform development time dropped from 90 days to 15 days **Lower release cost for internal IS** – The time to release for the internal test management system dropped from 37 man-days to two **Positive feedback from employees** – 95 percent of those asked would not want to return to the old way of working (pre-SAFe)

Costs & Benefits in euros

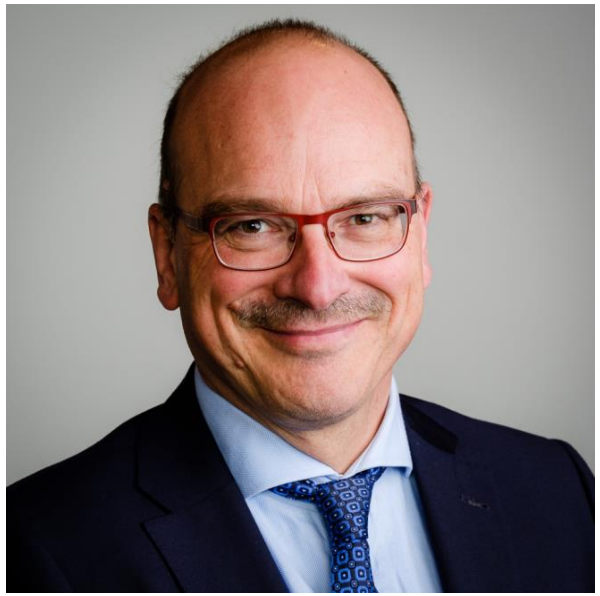
- No figures available yet from our own projects
- No figures available on the internet
- Companies don’t like to share this kind of data
- Special research must be set up for this

Conclusions

- Testing in production is a powerful tool but an extension to the test process
- Necessary in the modern development methods. We have to speed up delivering software
- But think before acting. It could be very risky
- Not applicable to all areas
- Fast feedback loop of a product into the market
- Extra assurance before using a product on a large scale

Future work

- Determine more situations to apply testing in production
- Gain more experience
- Identify standard risks
- Estimate the costs of testing in production in relation to testing during development
- Collect metrics about the benefits of testing in production
- Develop standard monitoring approach



Jos van Rooyen
E-mail: jos.van.rooyen@identify.nl
Mobile: +31654906282
LinkedIn: Jos van Rooyen