Challenges in Cutting Edge CI

Real-life story

Gil Hoffer
@gilhoffer

Ravello Systems
http://www.ravellosystems.com
@ravellosystems

July 16, 2014

#jenkinsconf

Hello
I am

Gil

ravello systems

Applications | Library | Users | Log | Upgrade Account | Show me how | Help & Support | Account | ravello

All | Blueprints | Application: jenkins-flow-manageme...

Data is saved | Properties | Save as Blueprint | Update | Auto stop in: Never | Published in: Amazon / Virginia | 0 cloud errors | 0 design errors

VM Library | Edit | Save VM to Library | Search... | 7 VMs 4 Groups

Application | Network

Import VM

Server

(drag to canvas to add)

- Ravello-ADServer
- Ravello-Guacamole Loc...
- ubuntu-12.04-serv... 0%
- Windows 2008 Server R...
- Guacamole Local Server
- Windows 2008 Server R...
- Windows 2008 Server R...
- Windows Server 2012-pr...
- Babushka-CServer
- precise-server-cloudimg-...

Frontend Tier (2) | App Tier (2) | Storage Tier (2) | Billing Tier (1)

CAS | VNC | DNS | Server | DB | Hornetq | UTS

VM Properties

System: 4 CPUs, 6 GB Memory

Configuration Management: Chef

Using Chef

Network Connections (1)

1. DHCP

(set by dhcp)

External Access: Public IP More...

Supplied Services (3)

1. ssh All IPs 22 External

2. pgadmin All IPs 4000 External

VM is Started

External Access:

db-jenkinsflowmanagemen-iwrje3ck.sr..

85.190.178.5

postgresql	5432
ssh	22
pgadmin	4000

SSH

Ask us

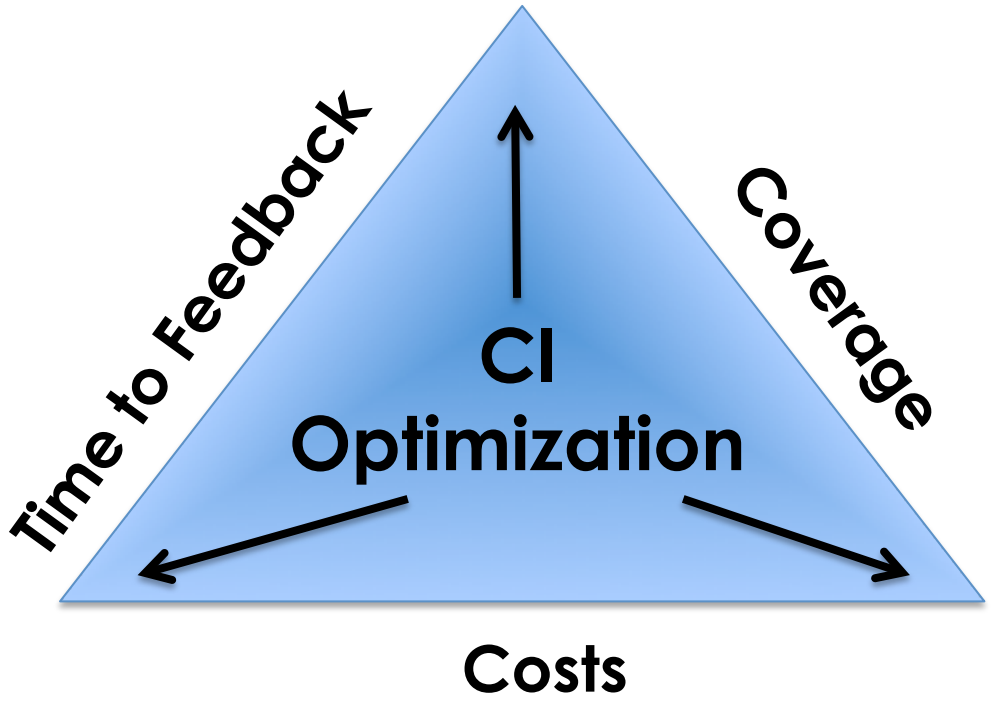
What are we going to talk about?

The CI 3-way tradeoff

"The Only Sure Thing in Computer Science – Everything is a tradeoff"

– Van Roy and Haridi.

Concepts, Techniques, and Models of Computer Programming



How it usually starts?
(if you're lucky)

- Unit tests, sometimes integration/system tests, automatic build – all are assumed for granted
- Someone setups a CI server
  - On-premise / cloud or service (cloudbees, TravisCI, …)
- We did have test automation infrastructure in mind from day one
- Test automation team?

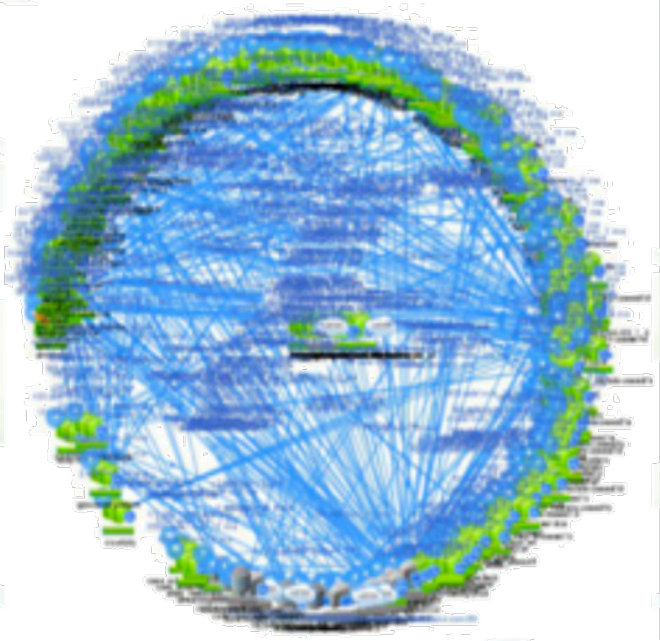
The team grows

- Managing a “process” with 10 engineers is >> 2x harder than managing a “process” with 5 engineers
- We grew from 5 to 22 in a year…

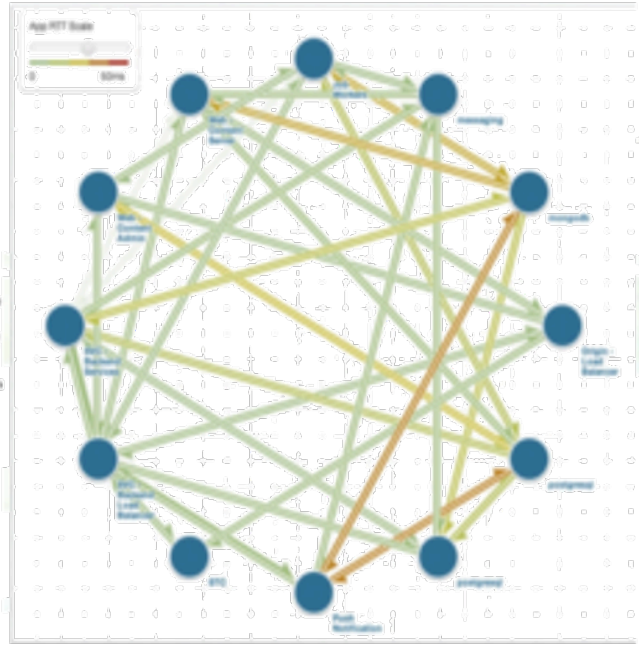
Complexity grows

original slide by Adrian Cockcroft

"Death Star" Architecture Diagrams



Netflix



Gilt Groupe (12 of 450)

Twitter

And…. Here come's GA…

From wikimedia commons (LGPL)

Complexity grows further

- Some more new components
- Data migrations
- Performance issues
- Zero Downtime upgrades
- Releasing our API
- …

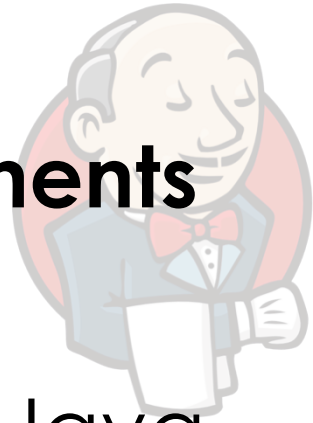
And we bump into “real life” issues

- Customers have issues
- Production incidents
- Security incidents
- Marketing / Product / Sales need that killer feature *yesterday*

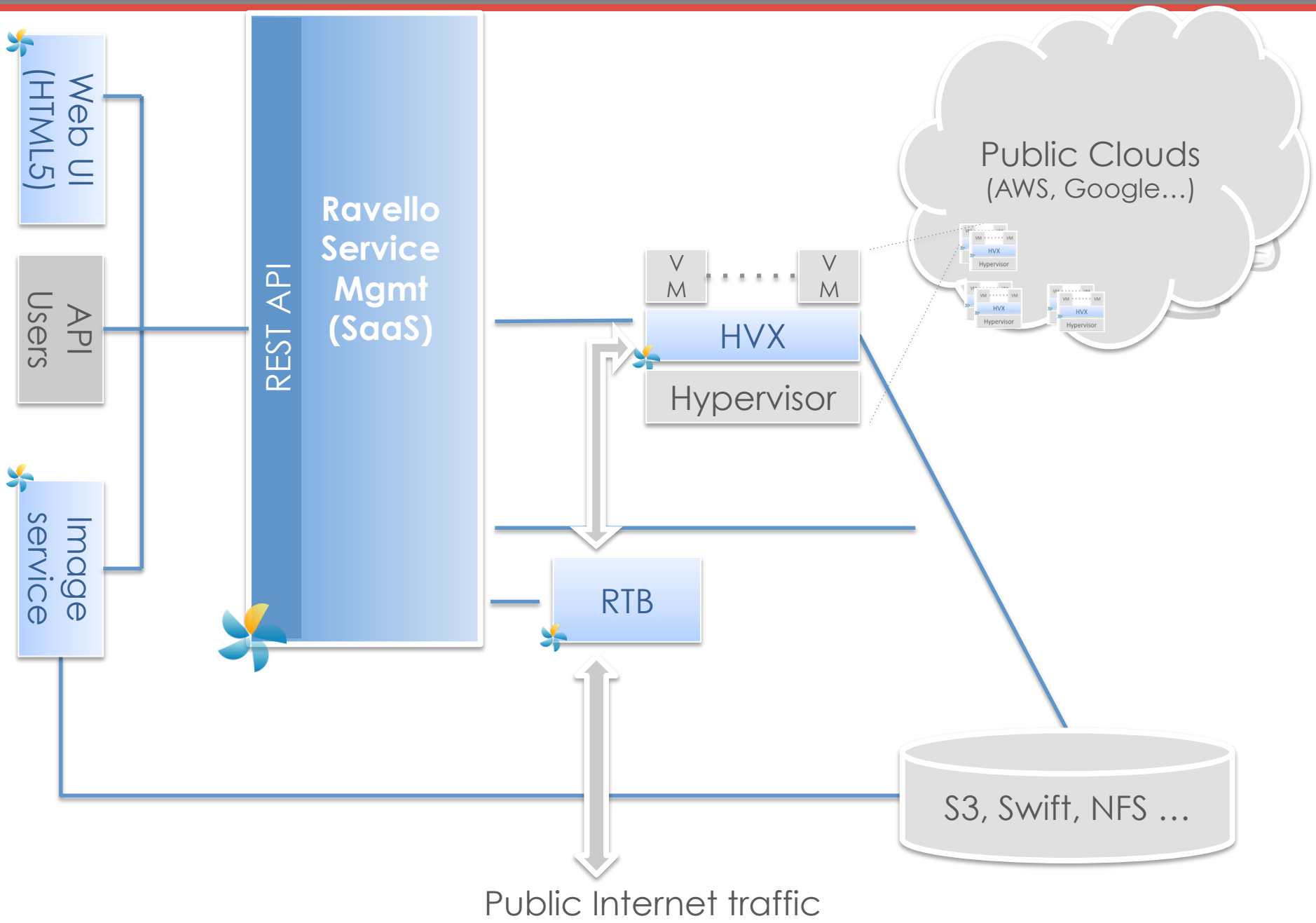
CI ?

- **C** for **continuous**
  - Yes, on every commit/push/check-in/…
- **I** for **integration**
  - Yes, you need to actually run **Integration** tests and not just unit tests
- If done properly (and that's a BIG if) can speed things up immensely
- If done poorly – can slow you down to a halt!

Context:
Ravello's core architecture components



- The HVX hypervisor (C, Python)
- Scale out management system (Java, JS, Cassandra, Postgres, Hazelcast, …)
- Cross platform upload utility (Python)
- Public internet traffic routing service
- Main external integration points:
  - Cloud providers (AWS, Google, Rackspace, …)
  - SaaSs (payment, CRM, marketing, emails, support , …)


Web UI (HTML5)
API Users
Image service
REST API
Ravello Service Mgmt (SaaS)
RTB
V M
V M
HVX
Hypervisor
Public Clouds (AWS, Google…)
S3, Swift, NFS …
Public Internet traffic

CHALLENGE ACCEPTED

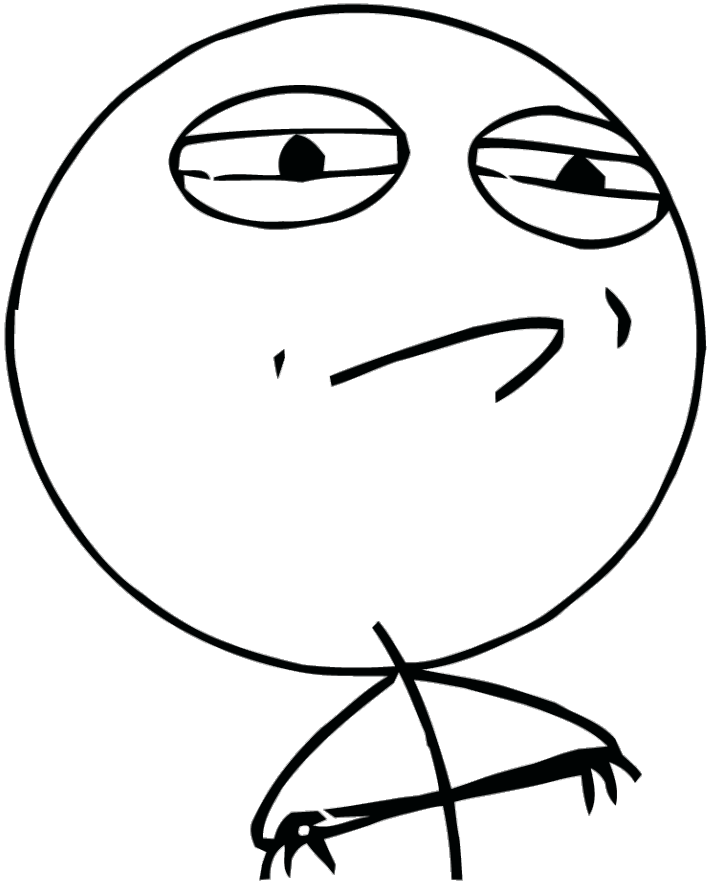
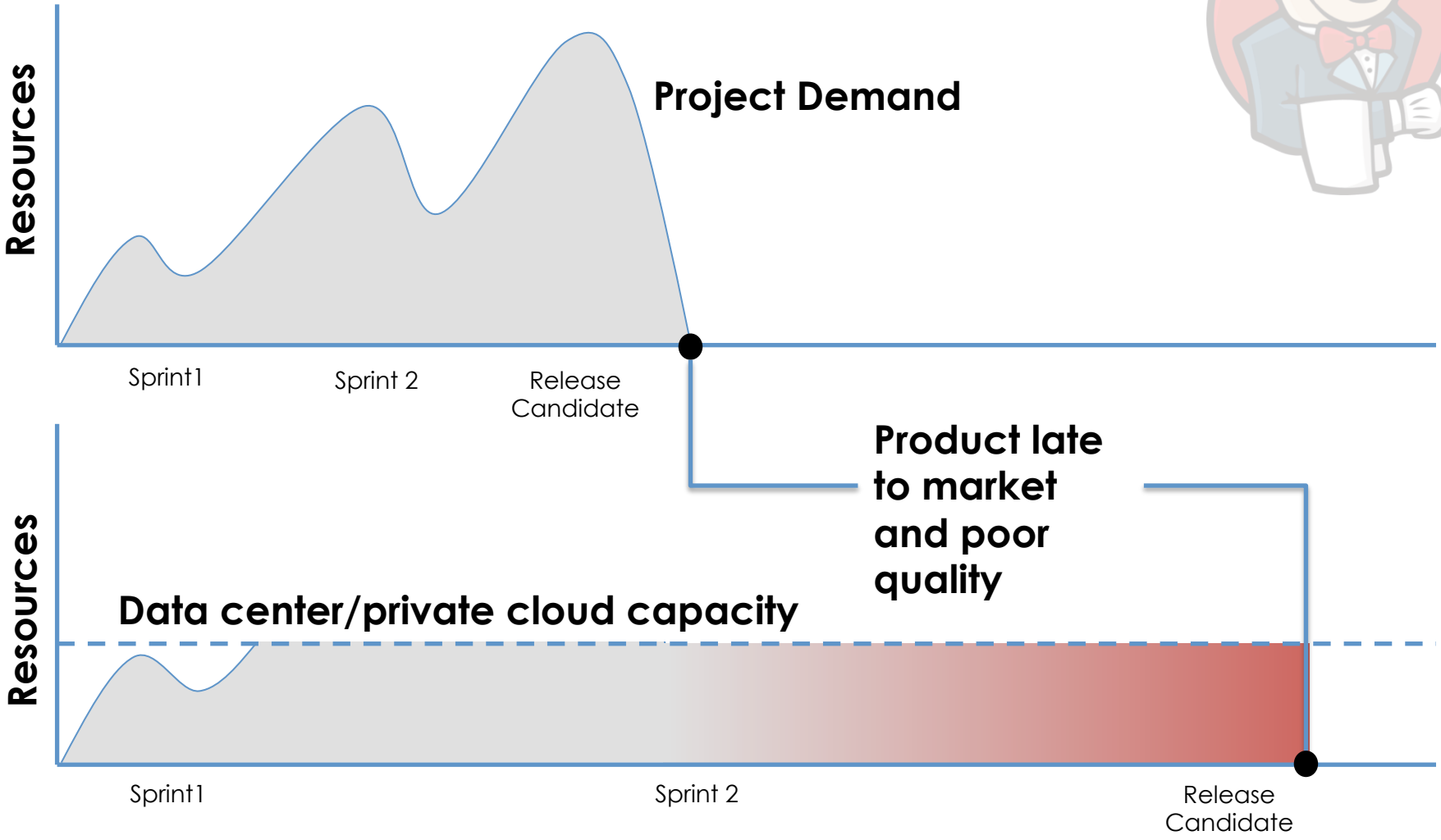


Image courtesy of alltheragefaces.com

Capacity


Resources
Project Demand
Sprint1
Sprint 2
Release Candidate
Product late to market and poor quality
Resources
Data center/private cloud capacity
Sprint1
Sprint 2
Release Candidate

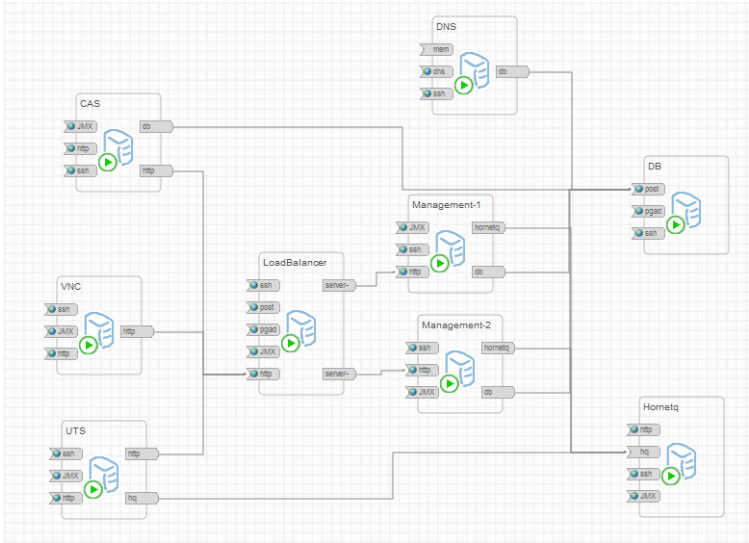
Capacity (2)

→ An “on-demand” consumption model for the bursty test/dev resources is A MUST

→ The cloud to the rescue!

Test Fidelity

- Test envs are as similar as possible to production
- We've crafted a solution which uses a combination of our own product (through a maven plugin), Jenkins and Chef to do so



Effective automatic UI testing

- Automatic UI testing is a controversial topic…
- We do believe that it is an important part of our CI pipeline
  - Test on Windows, Linux and Mac
  - Using Selenium Grid with nodes both in the cloud and on-premises (mostly for Mac (on Mac Mini))
  - Considered and disqualified Sauce-Labs due to cost

Full system testing

- We find full system testing to be **extremely** valuable
  - All components on LATEST
  - Single component on LATEST, all others on current PRODUCTION (stable) version
- In many cases, we do run partial / mocked / simulated components

Upgrade tests

- Upgrades of the system were one of the most painful items to test
  - Data migrations
  - Downtime implications
  - Resiliency
  - BW/FW compatibility issues

Upgrade tests

A three-phase approach

Production + Non Functional

- Make sure data migrations pass on real production data

A

Synthetic + Functional

- Synthetic functional upgrade tests
- Downtime implications

B

Production + Functional

- Keep and maintain long term test data on production
- Verify upgrade and functionality using it

C

Upgrade tests – deep dive (phase B)

- synthetic tests, each one with BEFORE and AFTER steps:
  - Install current PRODUCTION version
  - Do all BEFORE steps
  - Upgrade to version under test
  - Do all AFTER steps
  - *Assertions (also assert that operations that started BEFORE and were interrupted by the upgrade finished successfully)*

Feature Toggles

- Feature toggles are a must for any fast paced developed service
  - We've implemented toggles in every system component (UI, Backend, HVX, Uploader, …)
- BUT – how/when do you test them?
  - Our tests can turn on/off toggles
  - Only after some mileage in tests with a toggle on, we turn it on in production (gradually!)

Configuration Management and deployment mechanism tests

- It is CRUCIAL that your CI is using the same CM and deployment mechanisms as your production environments

- We run targeted CM unit tests, as we treat the CM as just yet another versioned component
  - We use Chef, Berkshelf and minitest/spec

Test Parameterizations

- The #1 requirement from your system test suites – to be able to be parameterized easily
- We heavily use TestNG's DataProviders to do so
  - Considering moving data to a test database from CSVs

Performance and scale tests

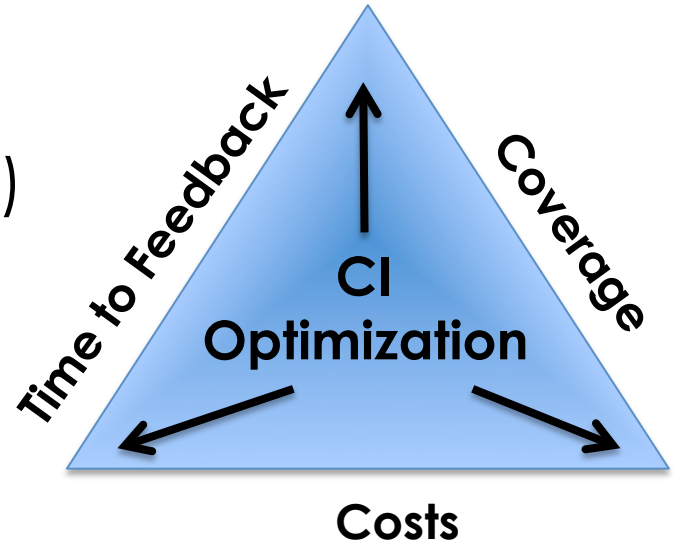
- We scaled up the service in some parameters much faster then expected
  - And we bumped into issues
- Performance tests should be a part of every automatic test process

False negatives are DEADLY

- False negatives are as worse as false positives if not worse
- Very quickly (a few days) deteriorate to a complete ignorance for the test results
- If you take one thing from this session:
  - **Exterminate false negative tests**

What to run when?

- **Smoke** (on each commit)
  - Build
  - UT
  - System smoke (all components!)
- **Sanity** (nightly)
  - System sanity
  - Upgrade A+B+C
- **Regression** (weekend)
  - System regression
  - Performance, stress, long runs



Something is broken…

- Fast analysis
- If you're responsible – you're on it…
- Stability on duty for all others
  - Need to have the ability to debug *fast*
    - Centralized logging (also for CI envs!)
    - Monitoring
    - Access to environments
- What about weekends?

Current State

ENABLE AUTO REFRESH

All | CI-Core-Smoke-Sanity | CI-Mgmt-Smoke-Sanity | CM-processes | DevOps-Processes | Manual-Process | Production-Rolls | Radiator-CI-Flows | Upgrade | Upload-Tool

S ↑	W	Name	Last Success	Last Failure	Last Duration
		management-gui-builder	46 min - #828	7 days 2 hr - #736	8 min 51 sec
		Babushka-provisioner	7 days 0 hr - #50	7 days 0 hr - #48	11 sec
		Chef-Server-GC	20 hr - #103	1 mo 18 days - #55	18 sec
		ci-resolve-lastgood-trunk-revision	2 min 21 sec - #898	23 days - #505	5.7 sec
		Cookbook-uploader	7 days 4 hr - #13	7 days 5 hr - #10	30 sec
		Dev-HVX-Builder-ubuntu1204	1 hr 10 min - #206	1 hr 12 min - #205	15 min
		devops-app-actions	10 min - #2868	4 days 4 hr - #2688	7 min 44 sec
		do-nothing	20 days - #12	N/A	61 ms
		flow-hvx-smoke	11 hr - #86	3 days 0 hr - #84	1 hr 51 min
		flow-publish-rmx-machine	15 hr - #108	3 days 2 hr - #107	6 min 27 sec
		flow-publish-tag	2 hr 35 min - #106	5 days 2 hr - #97	1 hr 21 min
		flow-publish-trunk	2 hr 29 min - #195	4 days 6 hr - #188	1 hr 46 min


flow-hvx-sanity
Now building: #69, 91% - 19 min left
Last Stable Build: 1 day 9 hr (in 3 hr 54 min)
flow-management-sanity
flow-management-smoke
Now building: #546, 15% - 1 hr 17 min left
Last Stable Build: 4 hr 58 min (in 1 hr 22 min)
Dev-HVX-Builder-ubuntu1204
flow-hvx-smoke
management-builder
Now building: #773, 2% - 18 min left

Jenkins

Jenkins ▸ Radiator-CI-Flows ▸ flow-management-smoke ▸ #538

- Back to Project
- **Status**
- Changes
- Console Output
- View Build Information
- Parameters
- Environment Variables
- Previous Build
- Next Build

Build #538 (Feb 20, 2014 8:01:06 PM)

No changes.

[BuildResultTrigger] A change to build result (log)

[Phase] - CREATE APPLICATIONS

devops-app-actions	build #2788	(7 min 42 sec)	parameters
devops-app-actions	build #2789	(8 min 1 sec)	parameters

[Phase] - CHECK NEW REVISION 1

ci-resolve-lastgood-trunk-revision	build #875	(8.7 sec)	parameters

[Phase] - BUILD UI

management-gui-builder	build #815	(9 min 45 sec)	parameters

[Phase] - CHECK NEW REVISION 2

ci-resolve-lastgood-trunk-revision	build #876	(8.5 sec)	parameters

[Phase] - CREATE WEBAPPS AND DEPLOY

flow-management-runtime-w-logs	build #470	(26 min)	parameters
flow-management-runtime-w-logs	build #471	(24 min)	parameters

[Phase] - RUN TESTS

automation-server-tests	build #793	(21 min)	parameters
automation-ui-tests	build #338	(10 min)	parameters
automation-server-tests	build #794	(17 min)	parameters

[Phase] - STOP APPLICATIONS

devops-app-actions	build #2792	(23 sec)	parameters
devops-app-actions	build #2793	(16 sec)	parameters

[Phase] - DELETE OLD BLUEPRINT

devops-app-actions	build #2794	(21 sec)	parameters

[Phase] - CREATE BLUEPRINT FROM APPLICATION

devops-app-actions	build #2795	(12 sec)	parameters

[Phase] - DELETE APPLICATIONS

devops-app-actions	build #2796	(15 sec)	parameters

Tools we're using

- Jenkins
  - Multi Job plugin
- TestNG
- Selenium Grid
- Artifactory
- Ravello
- Chef, Berkshelf, minitest/spec
- Vagrant
- Phornix test suites

Company wide impacts

- Responsibility
- Responsibility
- Responsibility
- Who writes and maintains system tests?
- The position of a test automation team?
- The role of manual QA

Conclusions

- The entire company must be bought into the process, it is *not* the business of just the test automation team or RnD
- Continuous integration != Test automation
- Continuous integration != Continuous delivery
- Always prefer evolution over revolution, but don't be afraid to make revolutionary changes if needed
- Visibility is crucial, false negatives are deadly

Questions?

Thank You To Our Sponsors

Platinum

Gold

Silver