



#103 - Tuesday, 10:30am

**A to Z Testing in Production:
Industry Leading Techniques
to Leverage Big Data for Quality**

Seth Eliot
Senior Knowledge Engineer in Test



About Seth

present

Microsoft Engineering Excellence

- Best practices for services and cloud



- Bing “Cosmos”
 - Massive, distributed, data processing service
- Microsoft Experimentation Platform
 - Data Driven Decision Making



Amazon.com Digital Media

- Video, Music, and Kindle services

past

TiP A-Z

A	Active Monitoring	N	Not to Do, What
B	Big Data	O	Operations
C	Client Instrumentation	P	Passive Monitoring
D	Data Mining	Q	Quality Signal
E	Experimentation	R	Real Data Input
F	Fault Injection	S	System Stress
G	Go / No-go for deployment	T	TestOps
H	High Availability	U	Users
I	Iterative Virtuous Cycle	V	3 V's of Big Data
J	JSI: J-script Instrumentation	W	Write Once, Test Anywhere
K	Killing production instances	X	eXposure Control
L	Load Testing in Production	Y	Y TiP?
M	Most Interesting Man in the World	Z	Zymurgy

Introduction

• • •

Testing in Production (TiP)

TestOps

Big Data

Testing at Microsoft 1985

- Design, execute and document tests
- Generate Test Scripts and automatic testing packages

programming

SOFTWARE TESTER

Microsoft's leadership in software dates to the advent of the microcomputer in 1974. With over two million installations, our first product, Microsoft BASIC, is the most widely used language in the world. As part of the Microsoft team, you will share our commitment to product development and technical excellence. We currently are seeking Software Testers to work on our newest applications software products.

As a Software Tester, you will design, execute and document tests of application software. You will generate test scripts and automatic test packages. This is a challenging and highly visible position within a fast growing division of Microsoft.

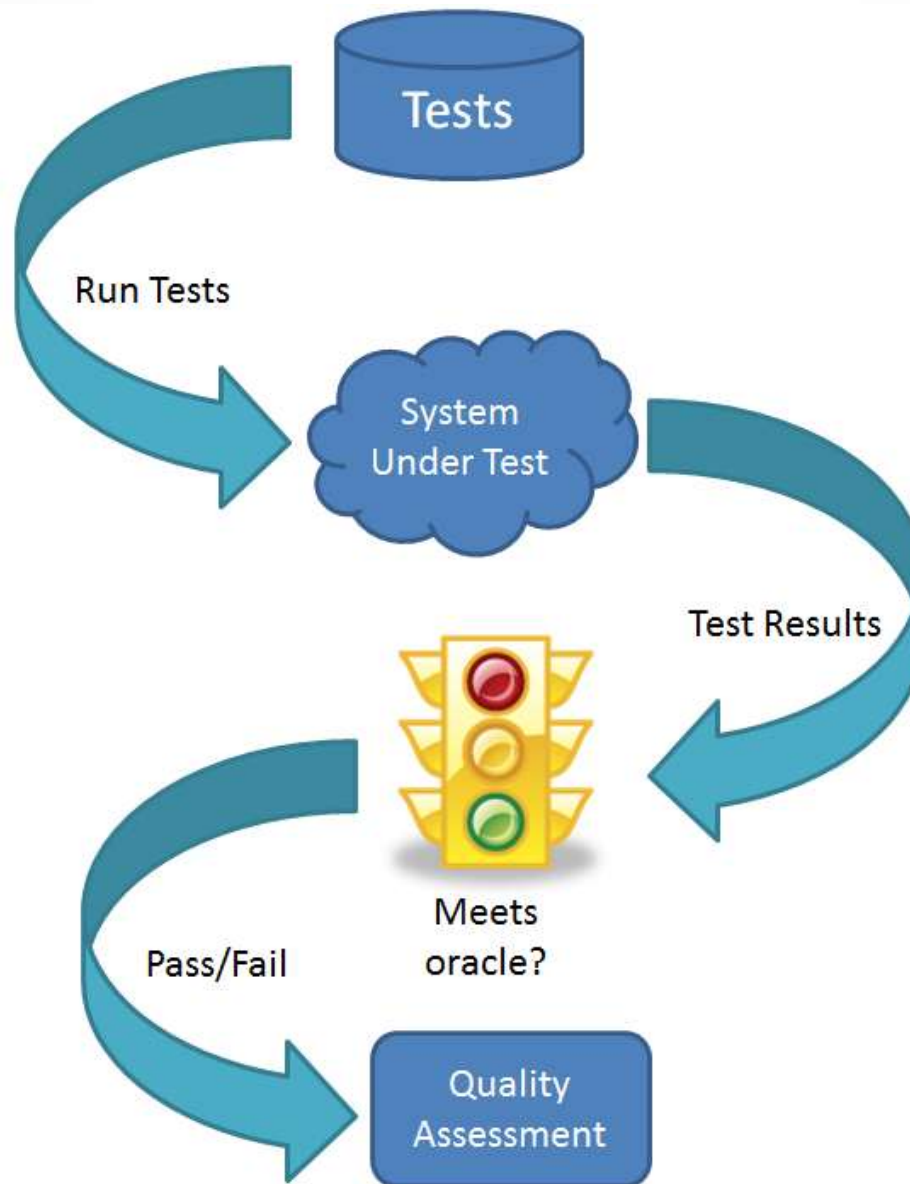
Requirements: background in math, computers, programming; preferably two years software test experience. Must be detail oriented, organized and have excellent written and oral skills.

Explore your future with Microsoft by sending a resume or letter that outlines your education, experience, and salary requirements to: Microsoft Corporation, Dept. ST, 40700 Northup Way, Box 97200, Bellevue, WA 98009. We are an equal opportunity employer.

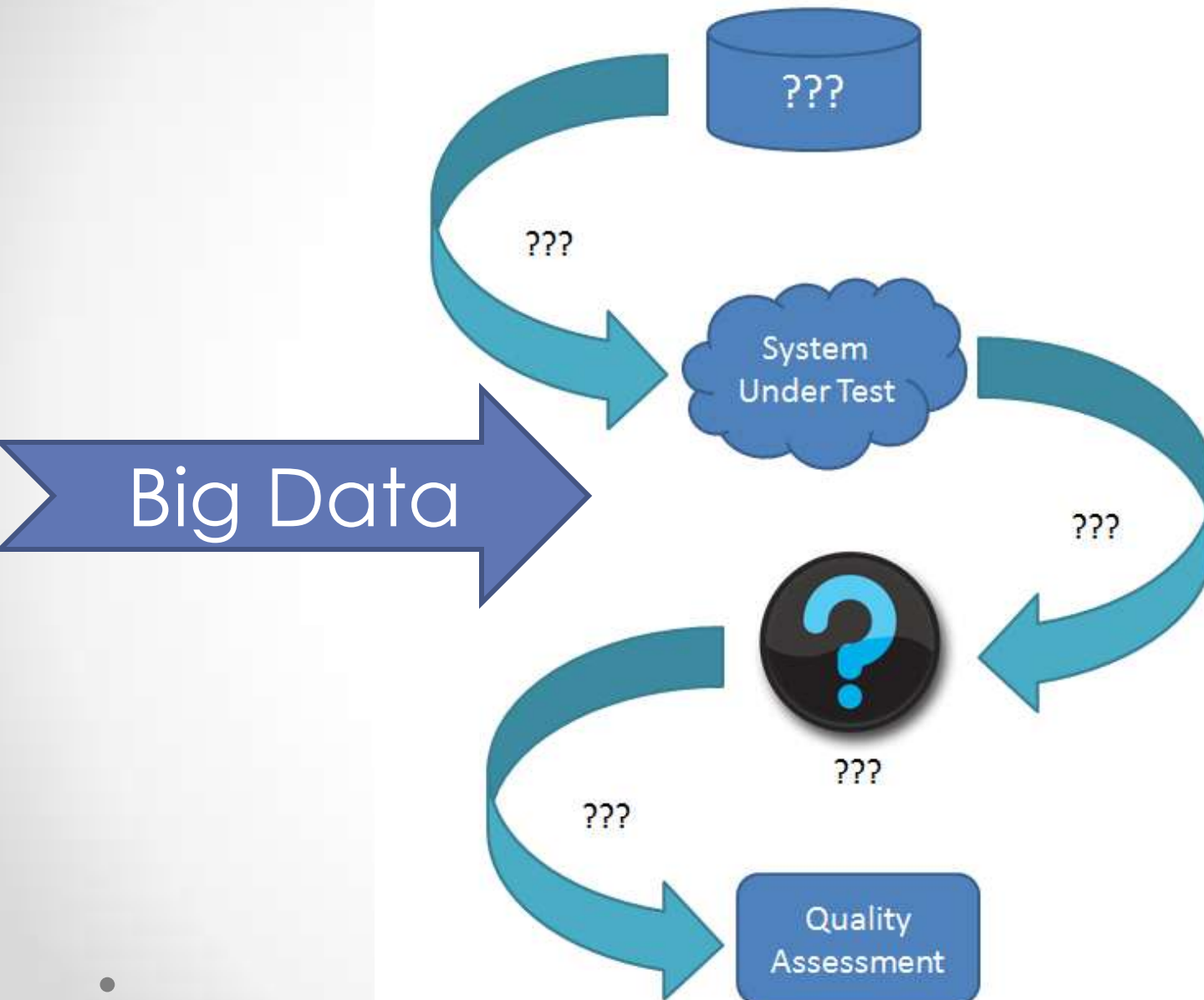
MICROSOFT

The High Performance Software™

What Testing Usually Is...



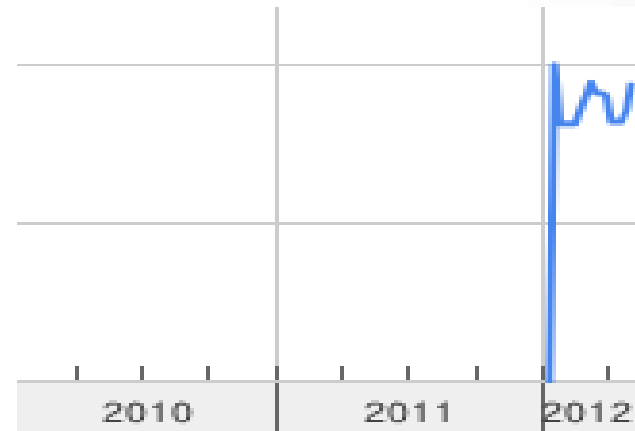
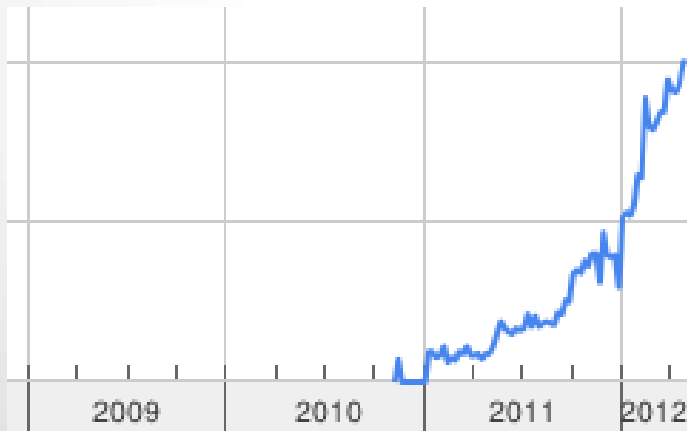
What Can Testing Be?



TestOps

- Monitoring: What Ops does
- Testing: What Test Does
- TestOps: Change the “signal” used for quality
From Test Results...

...to Big Data



The Three (or more) V's of Big Data

[Strata Jan 2012]

What is
Big
Data?

Volume

MB \Rightarrow GB \Rightarrow TB
PB \Rightarrow EB \Rightarrow ZB



Velocity



Variety



Value



The Big Data Signal

- Is often found in Production
- May not always be “Big”
 - The Quality Insights however should be Big
- **TestOps**: use this Big Data for quality assessment
- **Big Data** is in production
 - Therefore we **Test in Production**

WhY do we Test in Production?

- Leverage the diversity of real users
- ...and real prod environment...
- ...to find bugs you cannot find pre-production



The Most Interesting Man in the World



He's a Lover, not a fighter. But he's also a fighter so don't get any ideas

The police often question him, just because they find him interesting

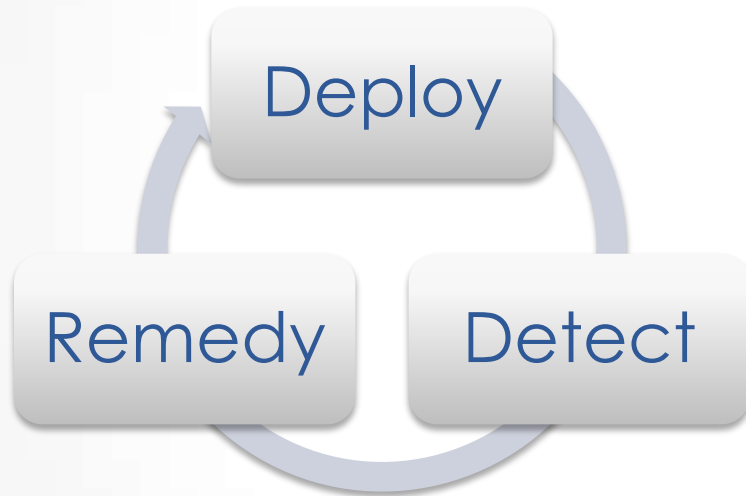
Will Testers Test in Production?



- “Stay buggy, my friends...”
- “That's our motto here. Doesn't work too well in practice, actually. “
- “Then blame all issues on QA -_- “
- “you don't even have to hire testers....they're just called ‘clients’.”

Why is TiP about Services?

- You control the deployment independent of user action.
- You have direct monitoring access.



- Iterative Virtuous Cycle

Google: All engineers have access to the production machines: “...deploy, configure, monitor, debug, and maintain”

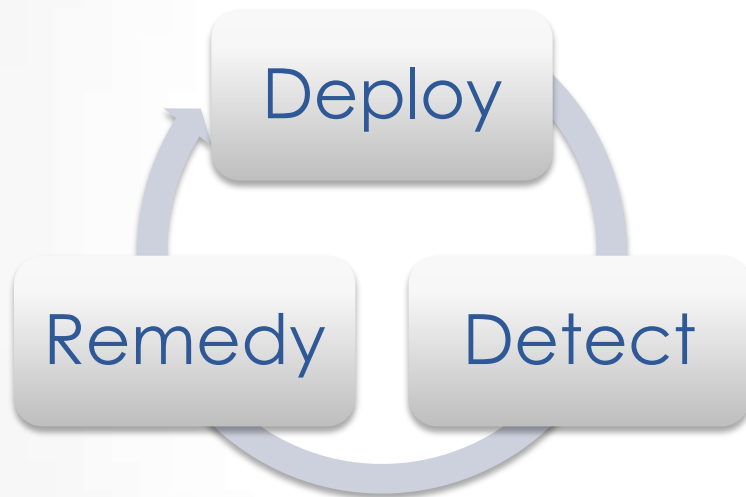
[Google Talk, June 2007 @ 21:00]

Facebook: engineers must be present in a specific IRC channel for “roll call” before the release begins or else suffer a public “shaming”

[Facebook ships, 2011]

Why is TiP (mostly) about Services?

- You control the deployment independent of user action.
- You have direct monitoring access.



- Iterative Virtuous Cycle

Google: All engineers have access to the production machines: “...deploy, configure, monitor, debug, and maintain”

[Google Talk, June 2007 @ 21:00]

Facebook: engineers must be present in a specific IRC channel for “roll call” before the release begins or else suffer a public “shaming”

[Facebook ships, 2011]



The Big Data Pipeline

- Facebook: Developers Instrument Everything
- Amazon: Central Monitoring
 - Add some config → Trending and Alerts
- Netflix: Custom libraries + AWS CloudWatch



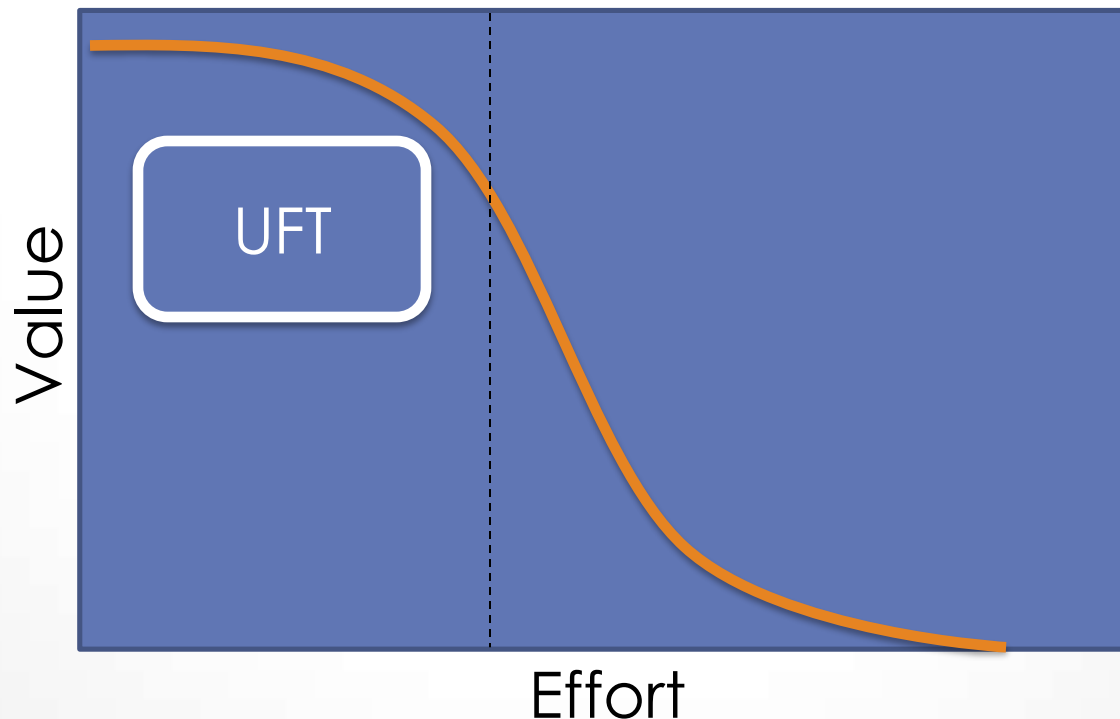
Servers

CPU

How does TiP fit into Test strategy?

Does TiP Replace Up-Front Testing (UFT)?

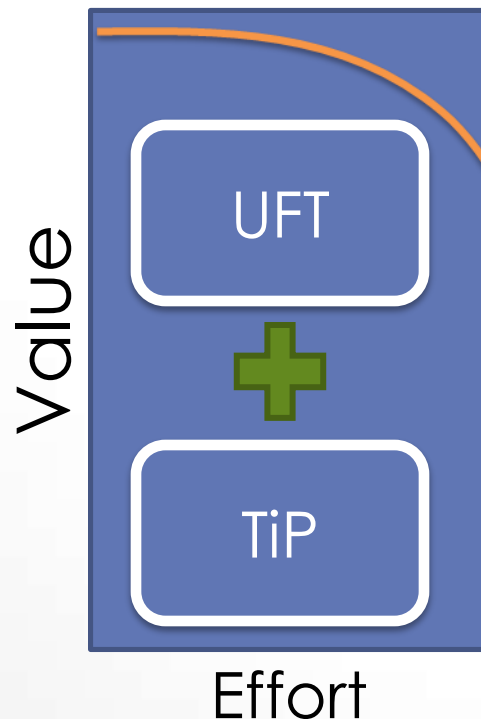
The Death of BUFT (Big UFT)?



How does TiP fit into Test strategy?

Does TiP Replace Up-Front Testing (UFT)?

The Death of BUFT (Big UFT)?



Four Categories of TiP

- Passive Monitoring

- with Real Data



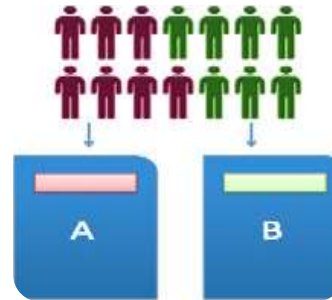
- Active Monitoring

- with Synthetic Transactions



- Experimentation

- on Real Users



- System Stress

- Of the Service and Environment



TiP A-Z

A	Active Monitoring	N	Not to Do, What
B	Big Data	O	Operations
C	Client Instrumentation	P	Passive Monitoring
D	Data Mining	Q	Quality Signal
E	Experimentation	R	Real Data Input
F	Fault Injection	S	System Stress
G	Go / No-go for deployment	T	TestOps
H	High Availability	U	Users
I	Iterative Virtuous Cycle	V	3 V's of Big Data
J	JSI: J-script Instrumentation	W	Write Once, Test Anywhere
K	Killing production instances	X	eXposure Control
L	Load Testing in Production	Y	Y TiP?
M	Most Interesting Man in the World	Z	Zymurgy

Passive Monitoring

...

with Real Data

Data Mining

- Bing stores and Analyzes Petabytes of data to
 - See real usage patterns
 - Find defects in quality
- Bing Relevance:
 - How can we test a complex algorithm without using that same algorithm
- Oracles often Heuristic
 - Find patterns for humans to follow-up



Data Mining: Speller Split Screen

The screenshot shows a Google search for 'fasfa'. The search bar at the top contains 'fasfa' and a magnifying glass icon. Below the search bar, there are tabs for 'Web', 'Videos', 'Images', and 'More'. The search results are displayed in two panes. The top pane, labeled 'Top Pane', shows the first two results. The first result is 'Home - FAFSA on the Web-Federal Student Aid' from 'www.fafsa.ed.gov'. The second result is 'Professional help filing the FAFSA - Federal Student Financial Aid ...' from 'https://www.fafsa.com'. The bottom pane, labeled 'Bottom Pane', shows the next three results. The first result is 'FASFA or FAFSA - Applications' from 'www.fasfa.com'. The second result is 'FAFSA - Free Application for Federal Student Aid' from 'https://fafsa.ed.gov'. The third result is 'FAFSA - Wikipedia, the free encyclopedia' from 'en.wikipedia.org'. The fourth result is 'FASFA' from an unknown source.

fasfa

Web Videos Images More

ALL RESULTS 1-11 of 455,000 results · [Advanced](#)

Were you looking for: [fafsa](#)

[Home - FAFSA on the Web-Federal Student Aid](#)
[www.fafsa.ed.gov](#)
Electronically submit the **Free Application for Federal Student Aid (FAFSA)**. All students interested in financial aid for college will need to complete this form.
[Contact Us](#) [Help](#)
[Español](#)

[Quick Access](#)
Customer service 800-433-3243

[Professional help filing the FAFSA - Federal Student Financial Aid ...](#)
Looking for assistance with the **FAFSA** application for federal student financial aid? We provide a wide range of services for a low fee.
[https://www.fafsa.com](#) · [Mark as spam](#)
[See more results](#)

Results for fasfa

[FASFA or FAFSA - Applications](#)
Looking for the **FASFA**? You may want FAFSA. We offer assistance with the FAFSA application for federal student financial aid at a low fee.
[www.fasfa.com](#) · [Mark as spam](#)

[FAFSA - Free Application for Federal Student Aid](#)
[https://fafsa.ed.gov/FOTWWWebApp/complete013.jsp](#) · [Cached page](#) · [Mark as spam](#)

[FAFSA - Wikipedia, the free encyclopedia](#)
[Eligibility](#) · [Types of financial aid](#) · [Filing options](#)
The Free Application for Federal Student Aid (known as the FAFSA) is a form that can be prepared annually by current and prospective college students (undergraduate and graduate) in ...
[en.wikipedia.org/wiki/FAFSA](#) · [Mark as spam](#)

[FASFA](#)
What is **FASFA**? The Frozen at Sea Fillets Association (**FASFA**) was formed in the autumn of 2000, to create awareness of the very high quality of Frozen at Sea Fillets of cod and ...

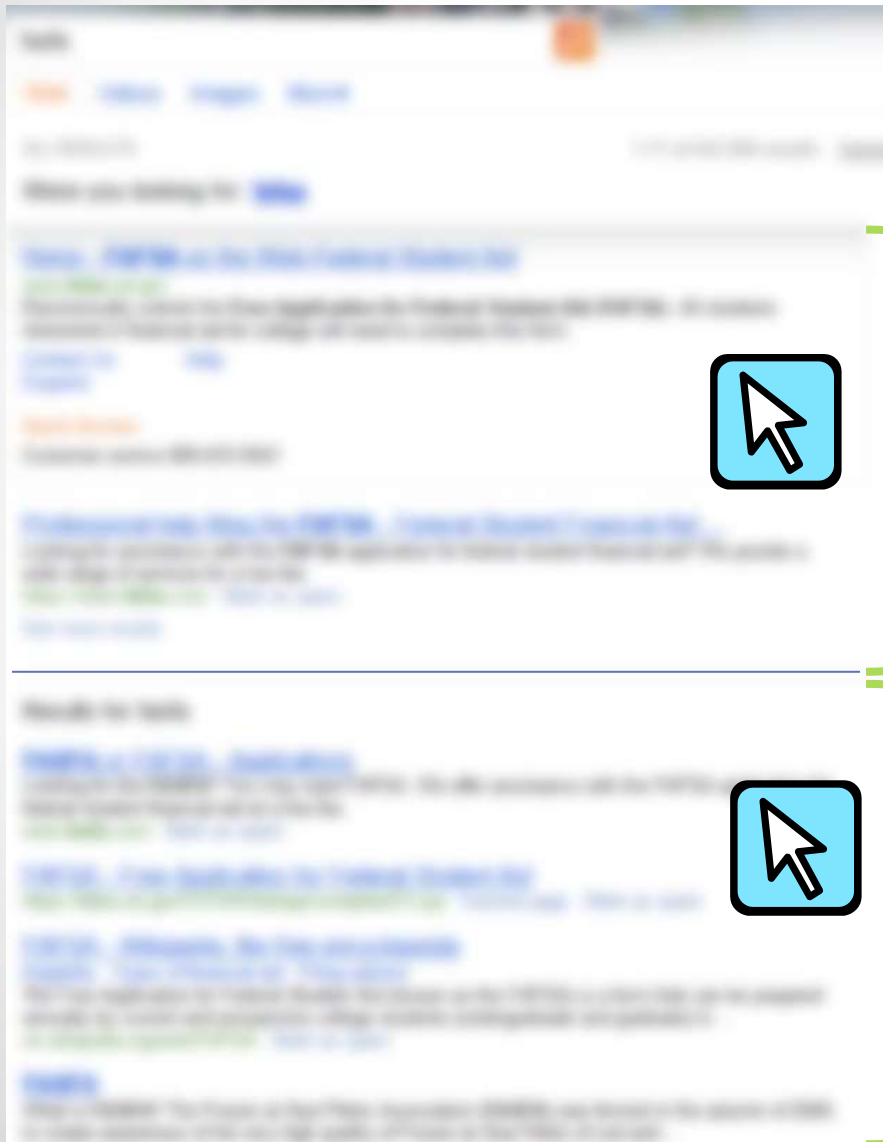
Top Pane

Result 1 and Result 2 from **Corrected** Query

Bottom Pane

Results from **Original** Query

Data Mining: Speller Split Screen



Alter query...

If top pane clicks > 90%

Leave query as-is...

If bottom pane clicks > 90%

Data Mining: Speller Split Screen

Some Top Results

Query	Correction	Top Clicks	Bottom Clicks	Keep Original	Make Correction
yah	yahoo	99.8%	0.2%		✓
fasfa	fafsa	90.5%	9.5%		✓
utube music	youtube music	90.4%	9.6%		✓
facebookcom	facebook.com	98.4%	1.6%		✓
imbd	imdb	96.3%	3.7%		✓
evony	ebony	0.5%	99.5%	✓	
century link	centrelink	2.0%	98.0%	✓	
yout	youth	3.8%	96.2%	✓	

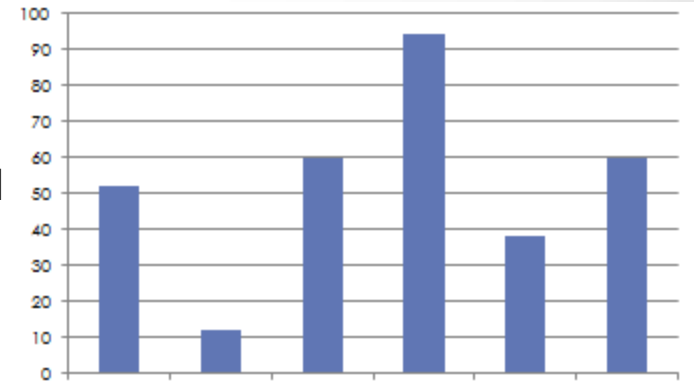
Detecting Video Takedown Pages



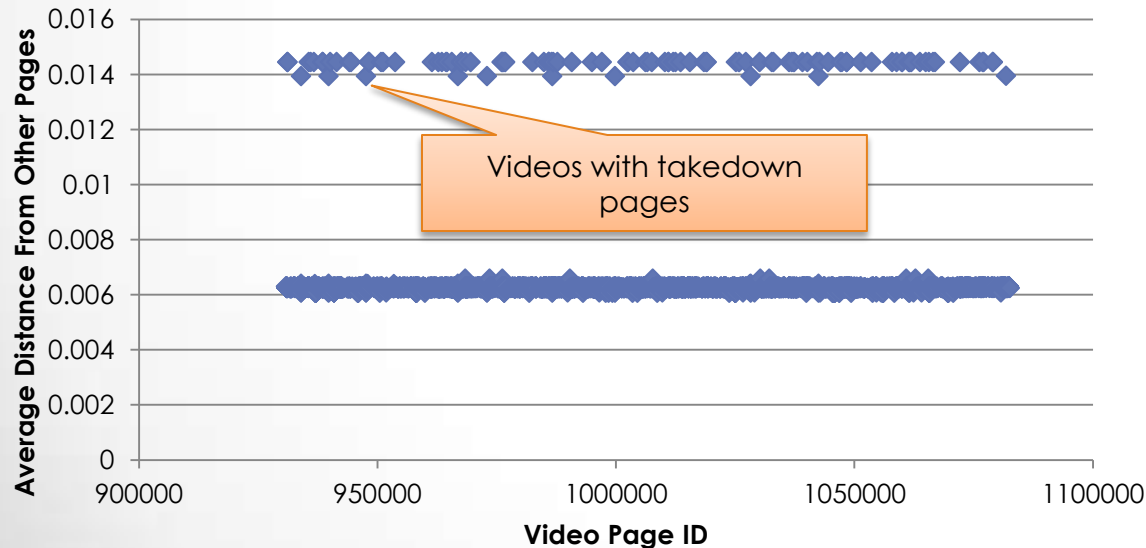
Automated heuristic:

Most videos on a video site will play. Takedown pages look different; find outliers.

Video Histogram "Fingerprint"



Average Page Distances - Fancast.com



- **5.56%** of our YouTube videos lead to a video takedown page.
- Thousands of users experience this per week.
- **We are feeding this back into the product!**

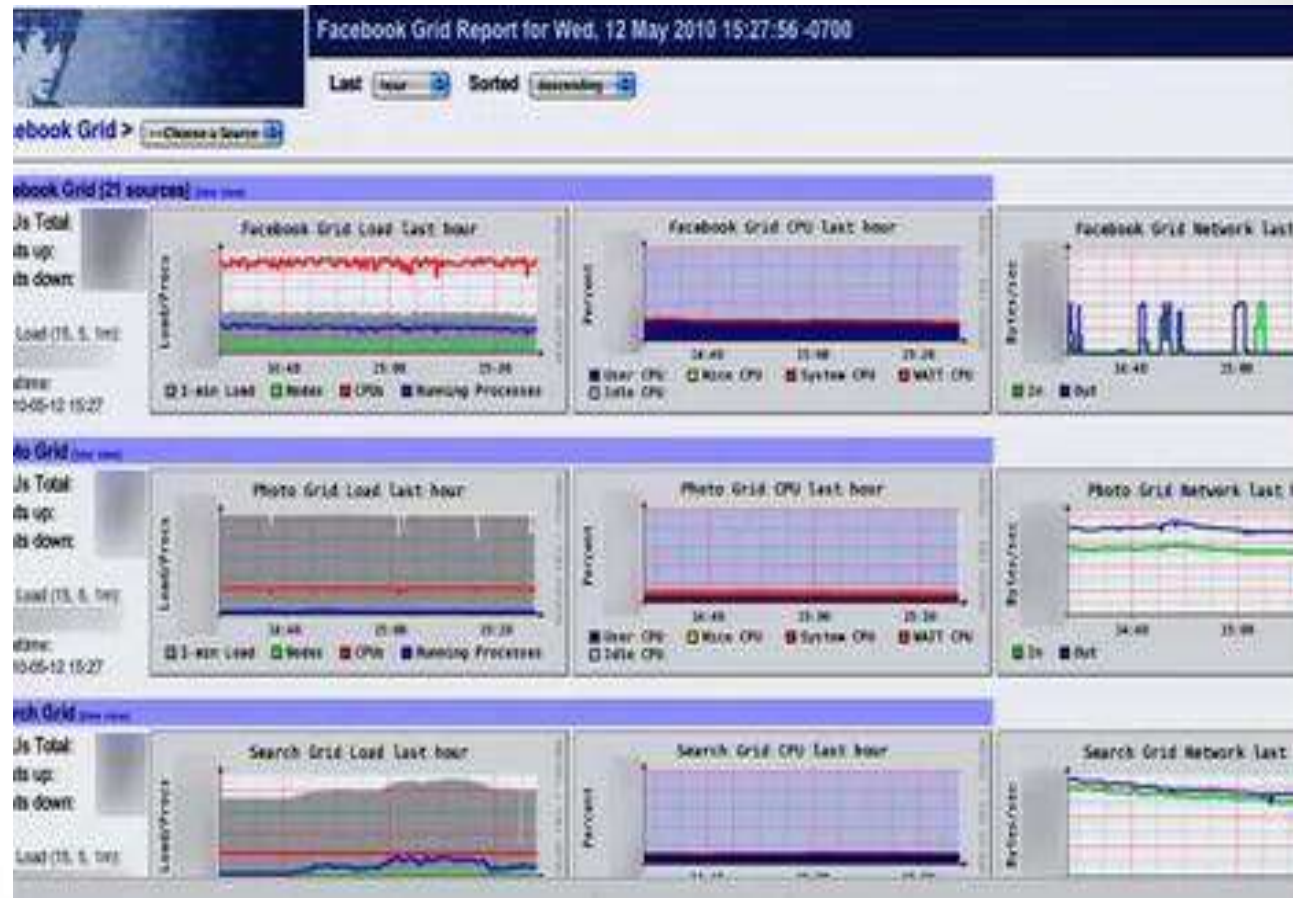


Facebook Mines Big Data for Quality



▶ “5 million metrics”

▶ CPU, network usage



User Performance Testing

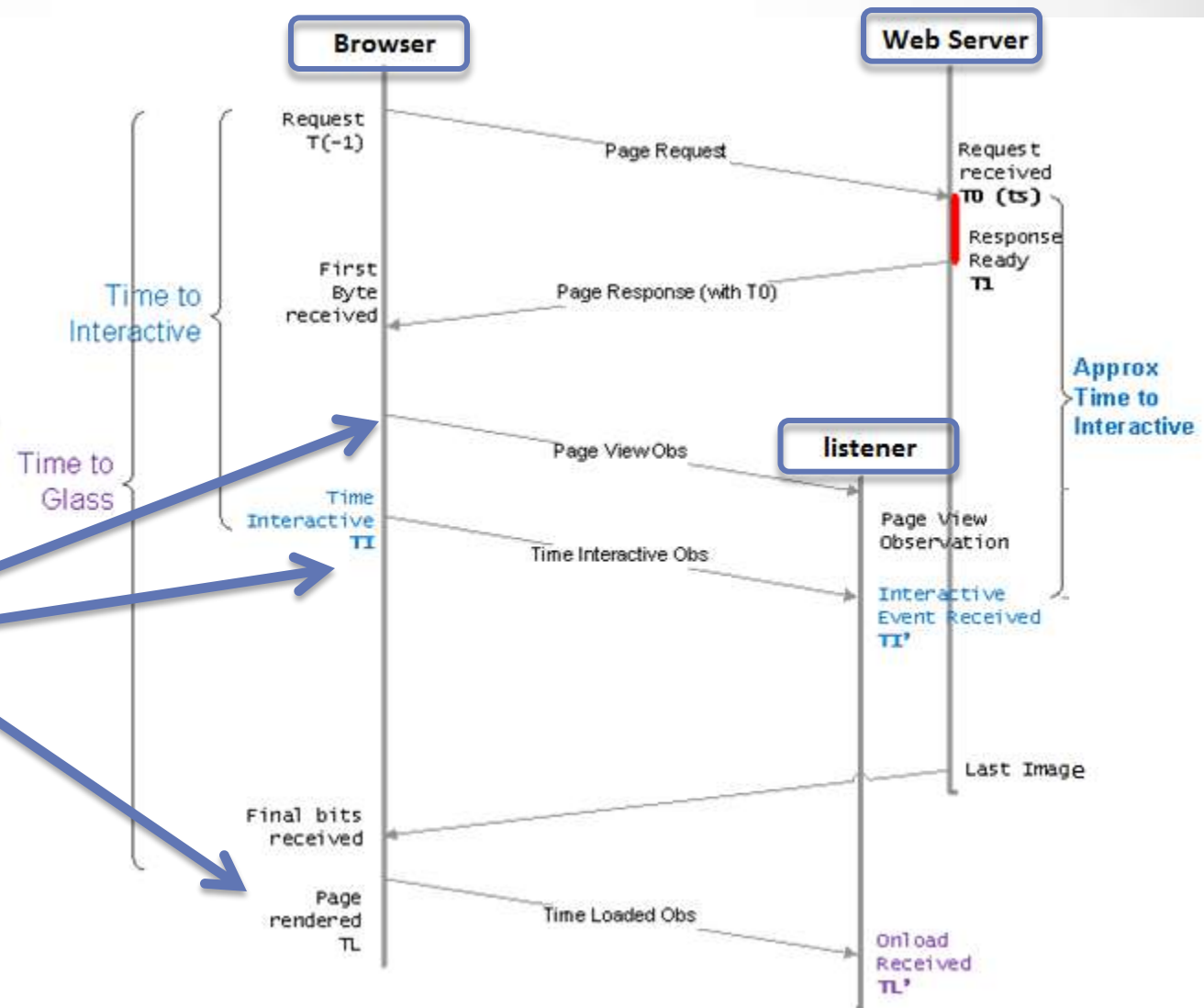
- Collect specific telemetry about how long **stuff** takes from user point of view
- Real User Data – Real User Experience
- End to End = complete request and response cycle
 - From user to back-end round-trip
 - Include traffic to partners, dependency response time
 - Measured from the user point of view
- From around the world
- From diversity of browsers, OS, devices



Stuff includes... Page Load Time (PLT)

TTG, real PLT
for real users

Client Instr.
sends
Beacons



Hotmail JSI User Performance Testing

- Customers do not care what we think our performance is - Aladdin Nassar, Sr. PM Hotmail

Big Data?

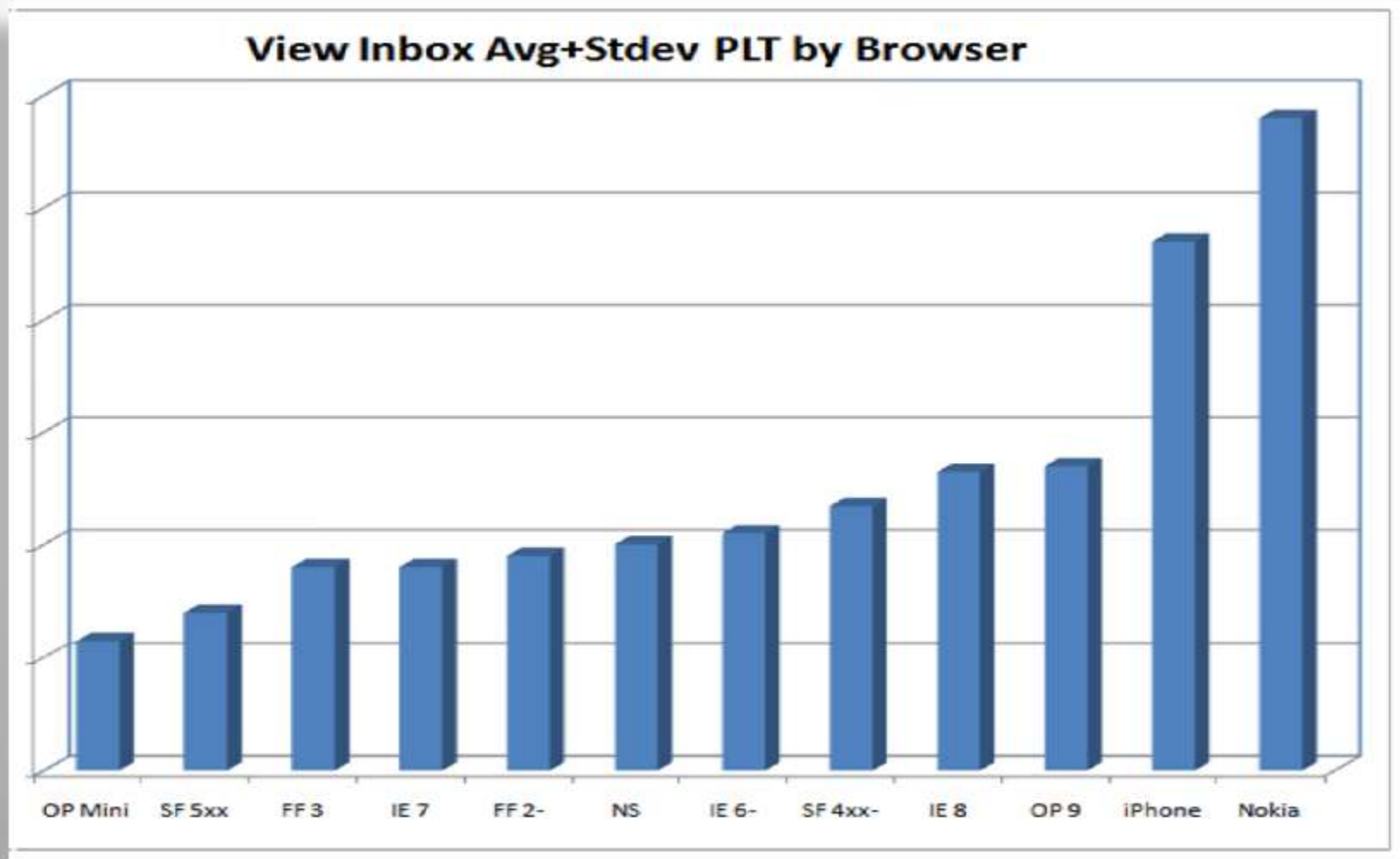
Hotmail's JavaScript Instrumentation (JSI)

- Budget for 500 Million measurements / month
- Scale for backend collection and analysis

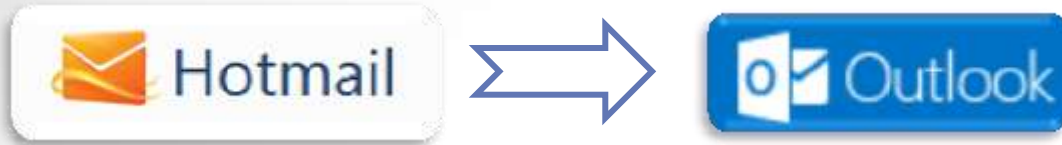
PLT by browser, OS, country, cluster, etc..

Hotmail JSI User Performance Testing

- PLT by browser, OS, country, cluster, etc..



User Performance Testing Examples



- Re-architected from the ground up around performance
 - Read messages are 50% faster
-
- Windows Azure™
 - **Every API:** Tracks how many calls were made; how many succeeded, and how long each call took to process

TiP A-Z

A	Active Monitoring	N	Not to Do, What
B	Big Data	O	Operations
C	Client Instrumentation	P	Passive Monitoring
D	Data Mining	Q	Quality Signal
E	Experimentation	R	Real Data Input
F	Fault Injection	S	System Stress
G	Go / No-go for deployment	T	TestOps
H	High Availability	U	Users
I	Iterative Virtuous Cycle	V	3 V's of Big Data
J	JSI: J-script Instrumentation	W	Write Once, Test Anywhere
K	Killing production instances	X	eXposure Control
L	Load Testing in Production	Y	Y TiP?
M	Most Interesting Man in the World	Z	Zymurgy

Active Monitoring

...

with Synthetic Transactions

TiP Test Execution

- From the Inside
 - Against internal APIs
 - Automated
- From the Outside
 - From User Entry Point
 - E2E Scenario in Production
 - Automated
 - or Manual

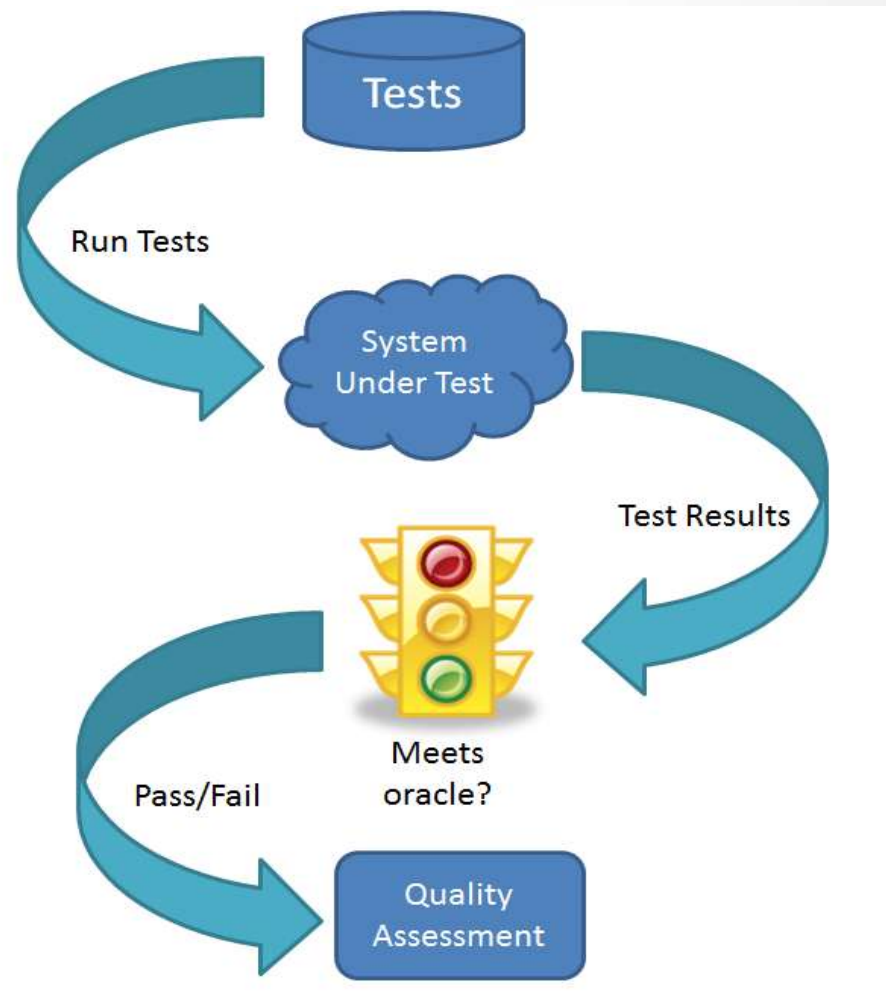


This looks like this

but in Production



which is OK, but...
Can we leverage
Big Data?



Active Monitoring

- Microsoft Exchange
 - Instead of pass/fail signal **look at thousands of continuous runs.**
 - Did we meet the "five nines" (99.999%) **availability** for scenario?
 - Is scenario slower this release than last? - **performance**

Scorecard Rows 142-177	Customer Experience					
Provisioning and ECP availability issue cause by partner load last month is fixed for current month	TIP Feature Availability					
IMAP log file issue from last month resolved and IMAP back to 100% availability	TIP Feature availability	Dec	Nov	Oct	Sep	Aug
New: Powershell scripts experiencing intermittent failure	Provisioning	99.93%	94.01%	100.00%	100.00%	100.00%
Improvements	RBAC	99.93%	98.81%	100.00%	100.00%	99.93%
• Content Delivery Network and DOMT tests added to address recent incidents	Outlook	100.00%	98.81%	100.00%	100.00%	99.95%
	ECP	100.00%	97.18%	100.00%	100.00%	99.97%
	Mailflow	100.00%	94.01%	100.00%	100.00%	99.50%
	ActiveSync	100.00%	99.73%	100.00%	100.00%	99.63%
	UM	100.00%	99.68%	100.00%	100.00%	99.68%
	Calendaring	100.00%	100.00%	100.00%	100.00%	100.00%
	Web Services	100.00%	100.00%	100.00%	100.00%	100.00%
	OWA	100.00%	100.00%	100.00%	100.00%	100.00%
	POP	100.00%	100.00%	100.00%	100.00%	100.00%
	IMAP	100.00%	89.87%	100.00%	100.00%	100.00%
	Remote PowerShell	99.93%	100.00%	100.00%	100.00%	100.00%



[Deschamps, Johnston, Jan 2012]

Test Data Handling

- Synthetic Tests + Real Data = Potential Trouble
 - Avoid it
 - Tag it
 - Clean it up

Example: Facebook Test Users

- Cannot interact with real users
- Can only friend other Test Users
- Create 100s
- Programmatic Control



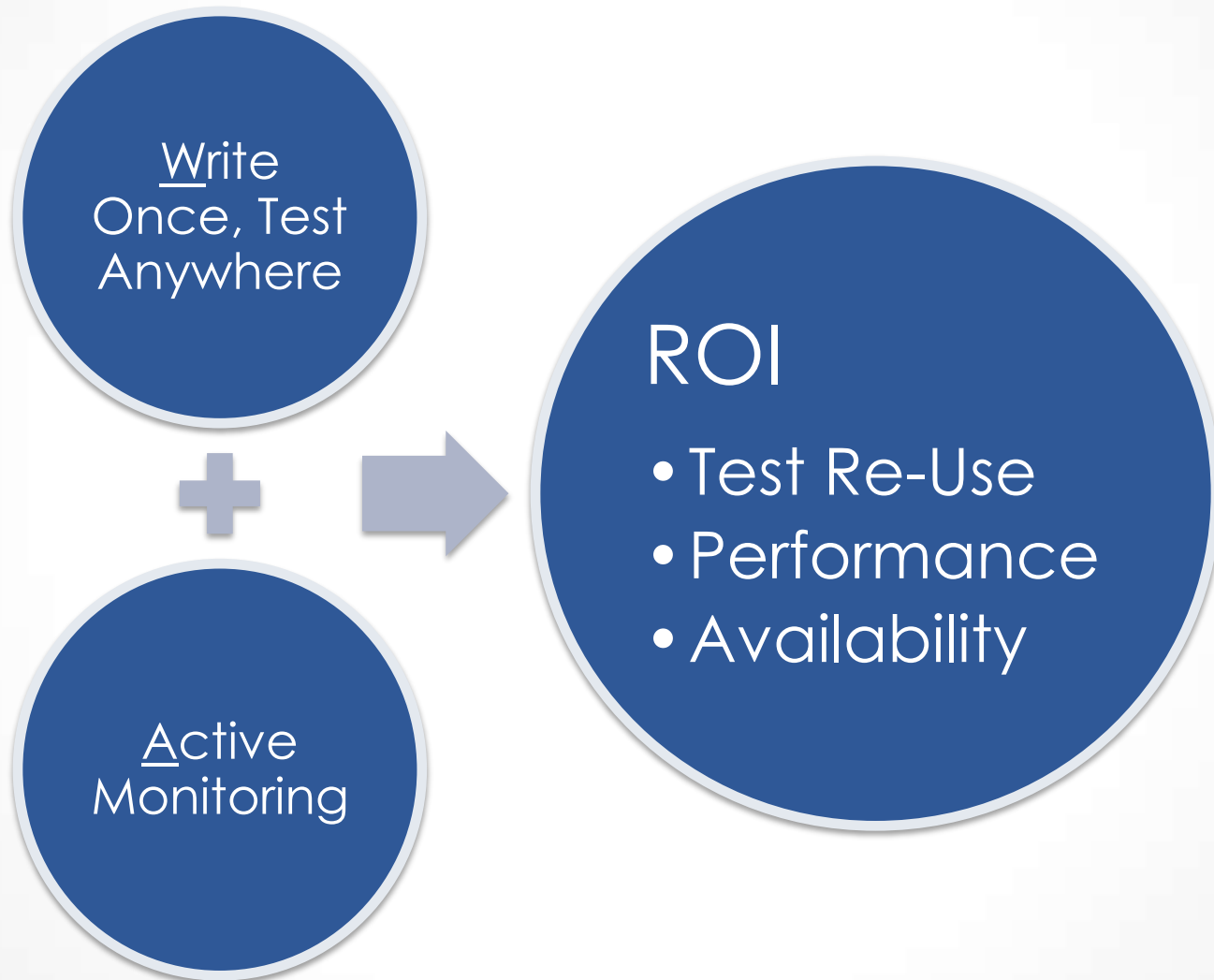
Write Once, Test Anywhere

- Microsoft Exchange
 - 70,000 automated test cases in lab
 - Re-engineered to run from the cloud
- TiP Execution Framework
 - Test Harness runs tests from Azure Cloud Platform
- Run key subset continuously against production



[Deschamps, Johnston,
Jan 2012]

More Testing, Less Cost



TiP A-Z

A	Active Monitoring	N	Not to Do, What
B	Big Data	O	Operations
C	Client Instrumentation	P	Passive Monitoring
D	Data Mining	Q	Quality Signal
E	Experimentation	R	Real Data Input
F	Fault Injection	S	System Stress
G	Go / No-go for deployment	T	TestOps
H	High Availability	U	Users
I	Iterative Virtuous Cycle	V	3 V's of Big Data
J	JSI: J-script Instrumentation	W	Write Once, Test Anywhere
K	Killing production instances	X	eXposure Control
L	Load Testing in Production	Y	Y TiP?
M	Most Interesting Man in the World	Z	Zymurgy

Experimentation

...

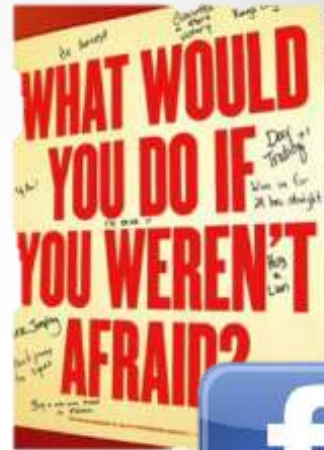
on Real Users



Experimentation

“To have a great idea,
have a lot of them”

-- Thomas Edison



- Try new things... in production
- Build on Successes
- Cut your losses... before they get expensive

Mitigate Risk with eXposure Control

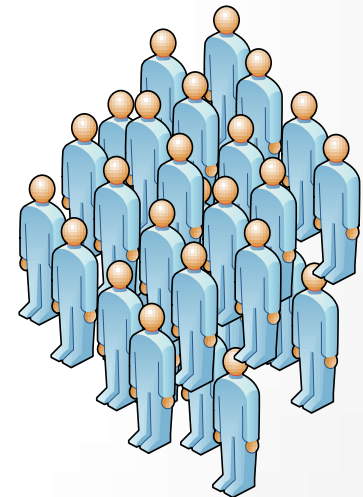
- Launch a new Service – Everyone sees it
- Exposure Control – only some see it



By Browser



By Location



By Percent
(scale)

Who's doing Exposure Control?

Three concentric push phases [Facebook ships, 2011]



- p1 = internal release
- p2 = small external release
- p3 = full external release

“We do these 1% launches where we float something out and measure that. We can dice and slice in any way you can possibly fathom.”

-Eric Schmidt, former CEO, Google

[Google BusinessWeek, April 2008]

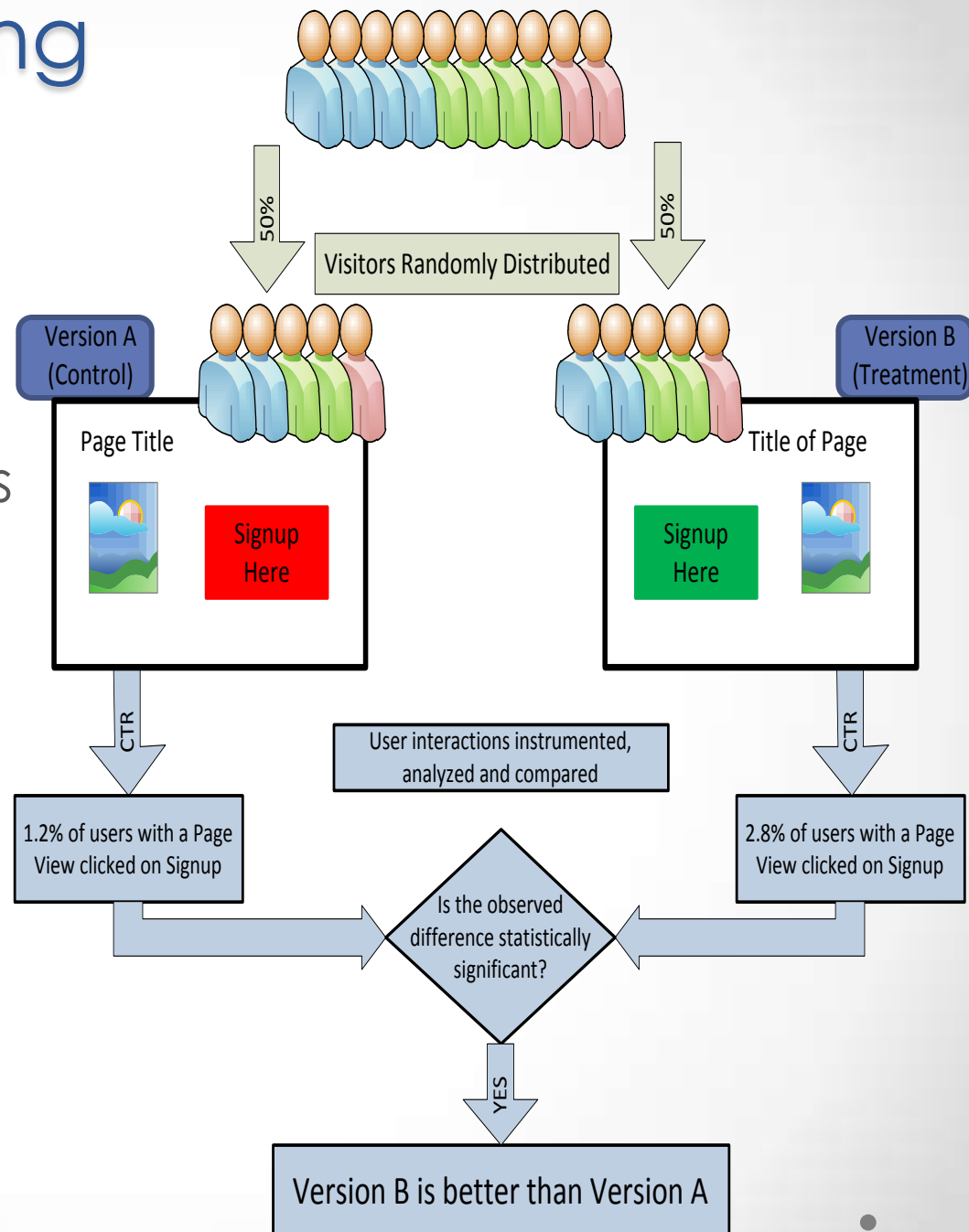


A/B Testing

- Aka: Controlled Online Experimentation

Big Data from Unique Uusers

	Expected Split	Actual Traffic
C: Control	12.50 %	17,420,461 unique users
T1	12.50 %	17,422,129 unique users
T2	12.50 %	17,414,591 unique users
T3	12.50 %	17,414,324 unique users
T4	12.50 %	17,418,765 unique users
T5	12.50 %	17,416,762 unique users
T6	12.50 %	17,417,993 unique users
T7	12.50 %	17,407,350 unique users




Experimentation for Design

Example: Microsoft Store

Goal: Increase Average Revenue per User

A

Shopping Cart

 Xbox 360 Elite Bundle \$399.99

Quantity: 1

[Remove](#) Delivery Options: Ship It

Shipping Method

☒ UPS Ground - \$0.00 (3-7 business days)
☐ UPS 2nd Day - \$5.00 (2 business days)
☐ UPS Next Day - \$12.00

Subtotal: \$399.99
Shipping: \$0.00
Tax: \$38.00


Estimated Total: \$437.99

Enter Zip to calculate Tax: 98052 [Calculate Tax](#)

[E-mail Cart to a Friend](#) [Checkout now](#)

B

Shopping Cart

 Xbox 360 Elite Bundle \$399.99

Quantity: 1

[Remove](#) Delivery Options: Ship It

Shipping Method

☒ UPS Ground - \$0.00 (3-7 business days)
☐ UPS 2nd Day - \$5.00 (2 business days)
☐ UPS Next Day - \$12.00

Subtotal: \$399.99
Shipping: \$0.00

[E-mail Cart to a Friend](#) [Checkout now](#)

Which increased revenue...?

A. Control

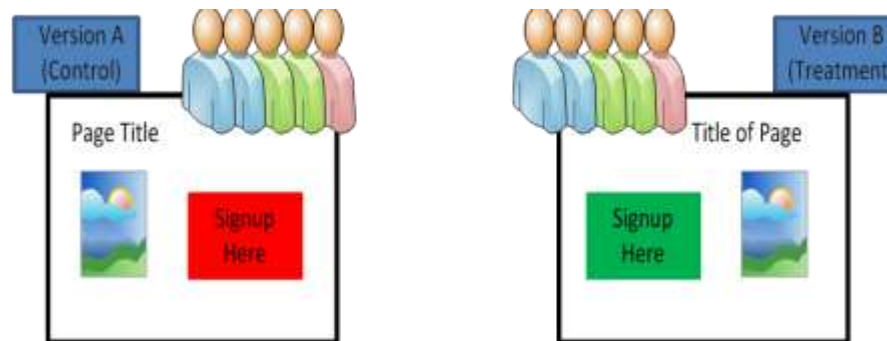
B. Treatment – up 3.3%

C. Neither

<http://store.microsoft.com/home.aspx>

AB Testing, Now for Code Quality

- Experimentation for Design
 - Business Metrics – Did we build the right thing?

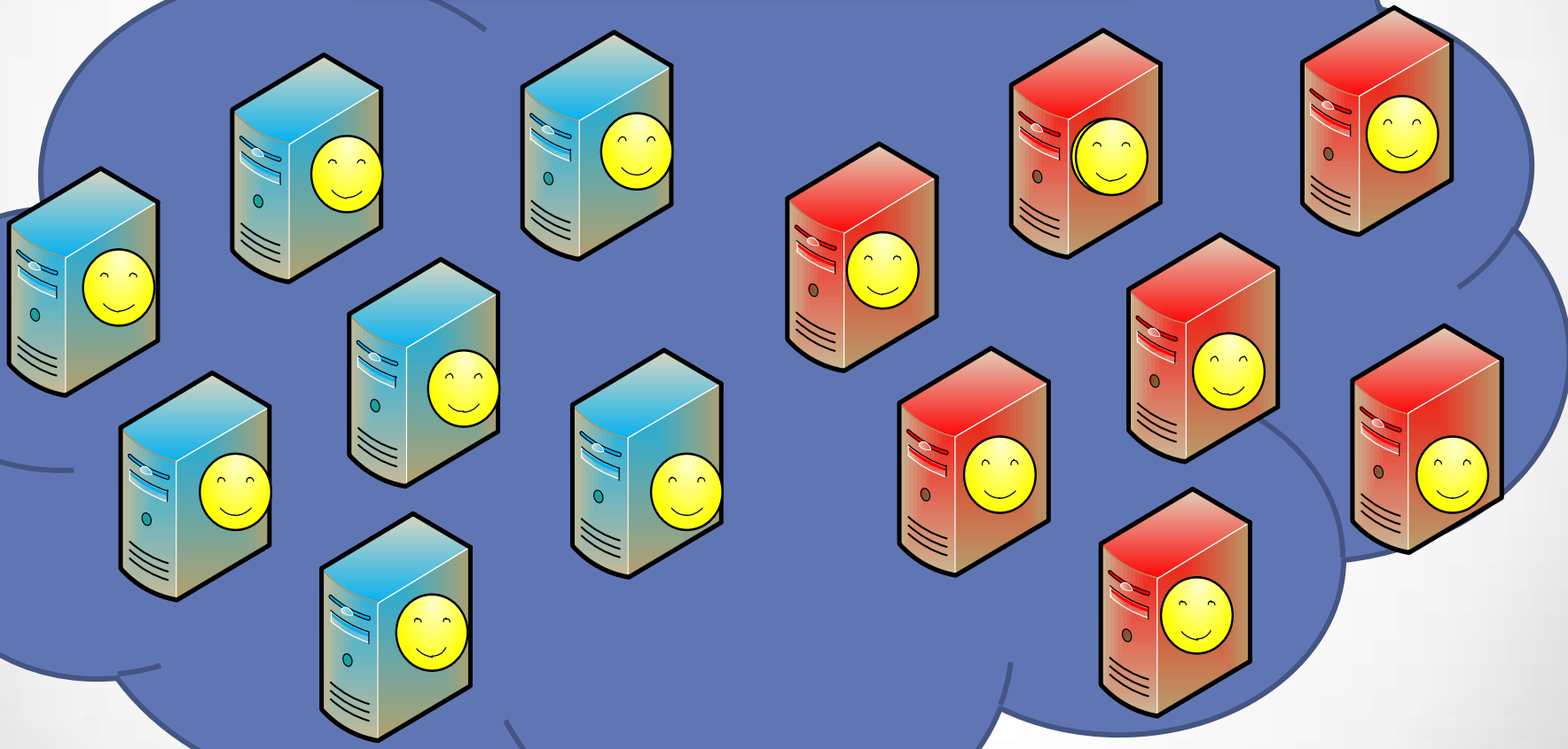


- Controlled Test Flight
 - Assess quality of new code → **deployment Go/No-go**
 - Code Quality Metrics –

Did we build it right?

Example: Controlled Test Flight: Netflix

1B API requests per day



“Canary” Deployment

Example Controlled Test Flight :

Amazon ordering pipeline

- Amazon's ordering pipeline (checkout) systems were migrated to a new platform.
- Team had tested and was going to launch.
- Quality advocates asked for a limited user test using Exposure Control.
- Five Launches and Five Experiments until $V_{curr} = V_{next}$ (showed no difference.)
- The cost had it launched initially to the 100% users could have easily been in the millions of dollars of lost orders.



Dogfood and Beta

The image shows the Bing homepage with a cave-themed background. A bright light beam from above illuminates a path in the cave. Two small figures are visible on the path. The Bing logo and search bar are in the top left. Navigation links (IMAGES, VIDEOS, MAPS, NEWS, SEARCH HISTORY, MORE, MSN, HOTMAIL) are at the top. A 'Sign in' link and user count '361' are in the top right. A black box with white text 'Un-Controlled Experimentation' is overlaid on the right. A 'Bing Feedback' panel is open on the bottom right, showing options to highlight or hide parts of the page, a text area for feedback, and a 'Submit' button. The bottom of the page features 'Popular now' (Hurricane Isaac path, Hantavirus, Rachel Corrie, Lindsay Lohan, GOP convention, Archbishop DUI), 'Search History' (Turn on search history to start remembering your searches), 'Images' (See all), and 'Videos' (See all). The footer includes copyright '© 2012 Microsoft' and links for Privacy and Cookies, Legal, Advertise, About our ads, Help, and Feedback. A 'Compare to Google' link is also present.

Un-Controlled Experimentation

Bing Feedback

Click to highlight parts of the page you want to comment on. Hide areas you want to keep private.

Highlight Hide

Tell us more

Problem Suggestion Kudos

Describe the problem, offer a suggestion, or let us know what we're doing right.

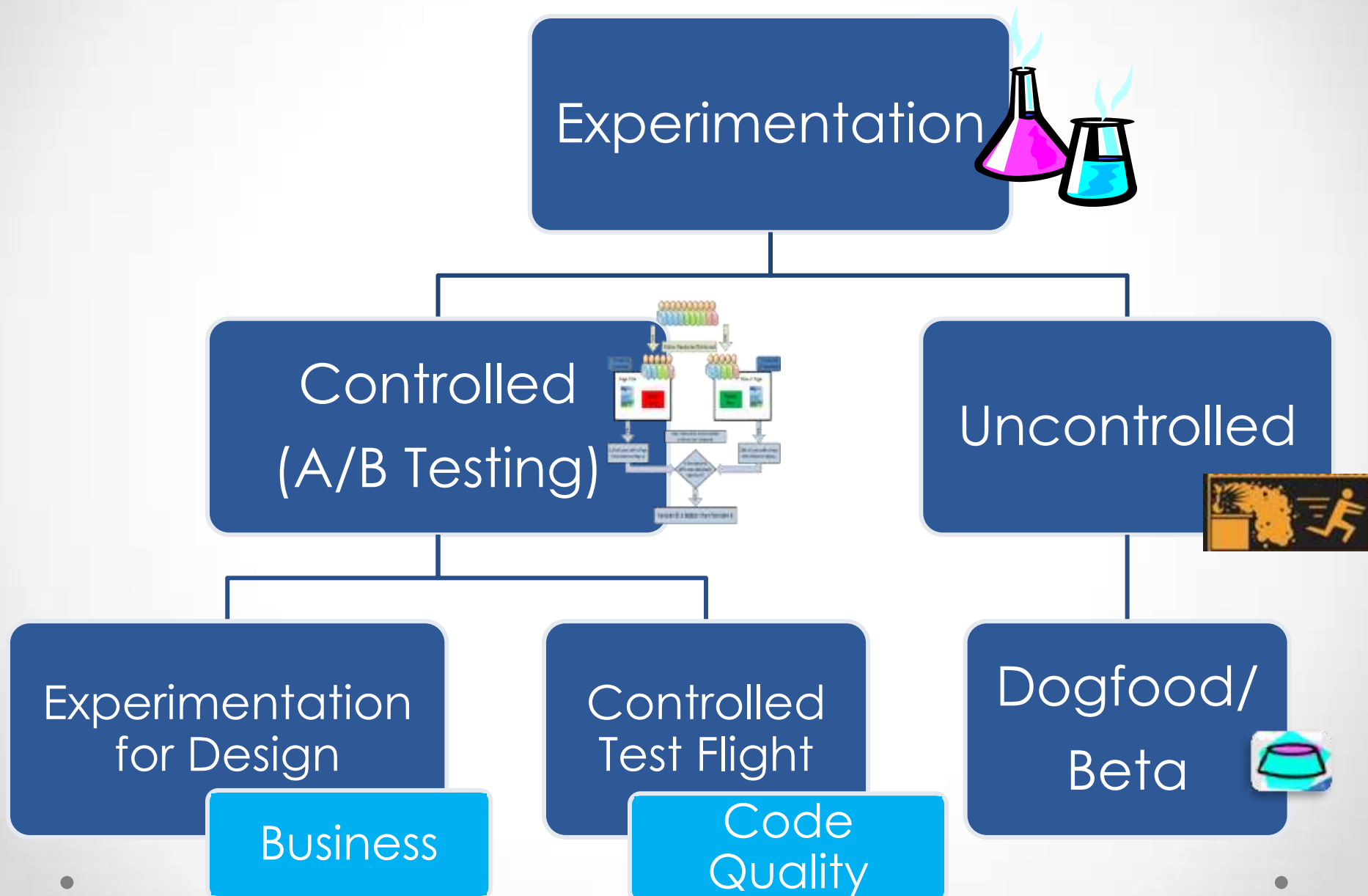
☒ Send email to the Bing team (sfeed) and cc me at your alias here

Submit

Compare to Google

Internal preview Help improve Bing

The Experimentation Family



TiP A-Z

A	Active Monitoring	N	Not to Do, What
B	Big Data	O	Operations
C	Client Instrumentation	P	Passive Monitoring
D	Data Mining	Q	Quality Signal
E	Experimentation	R	Real Data Input
F	Fault Injection	S	System Stress
G	Go / No-go for deployment	T	TestOps
H	High Availability	U	Users, Big-Data from Unique
I	Iterative Virtuous Cycle	V	3 V's of Big Data
J	JSI: J-script Instrumentation	W	Write Once, Test Anywhere
K	Killing production instances	X	eXposure Control
L	Load Testing in Production	Y	Y TiP?
M	Most Interesting Man in the World	Z	Zymurgy

System Stress

...

of the Service and Environment

Load Testing in Production

- Injects load on top of real user traffic
- Monitors for **performance**
- Load should **not** go through partners or data providers

Big Data

- Traffic mix: real user queries, simulate scenarios
- Real time telemetry: Monitor and Back-Off
- After the fact Analysis
 - Tune SLAs/Targets
 - Tune real-time monitors and alerts

Load Testing in Production

- Live Traffic + 1 million concurrent virtual users
- 77,000 hits per second
- 6 gigabits per second



- Identified major datacenter power issue

Effort went From

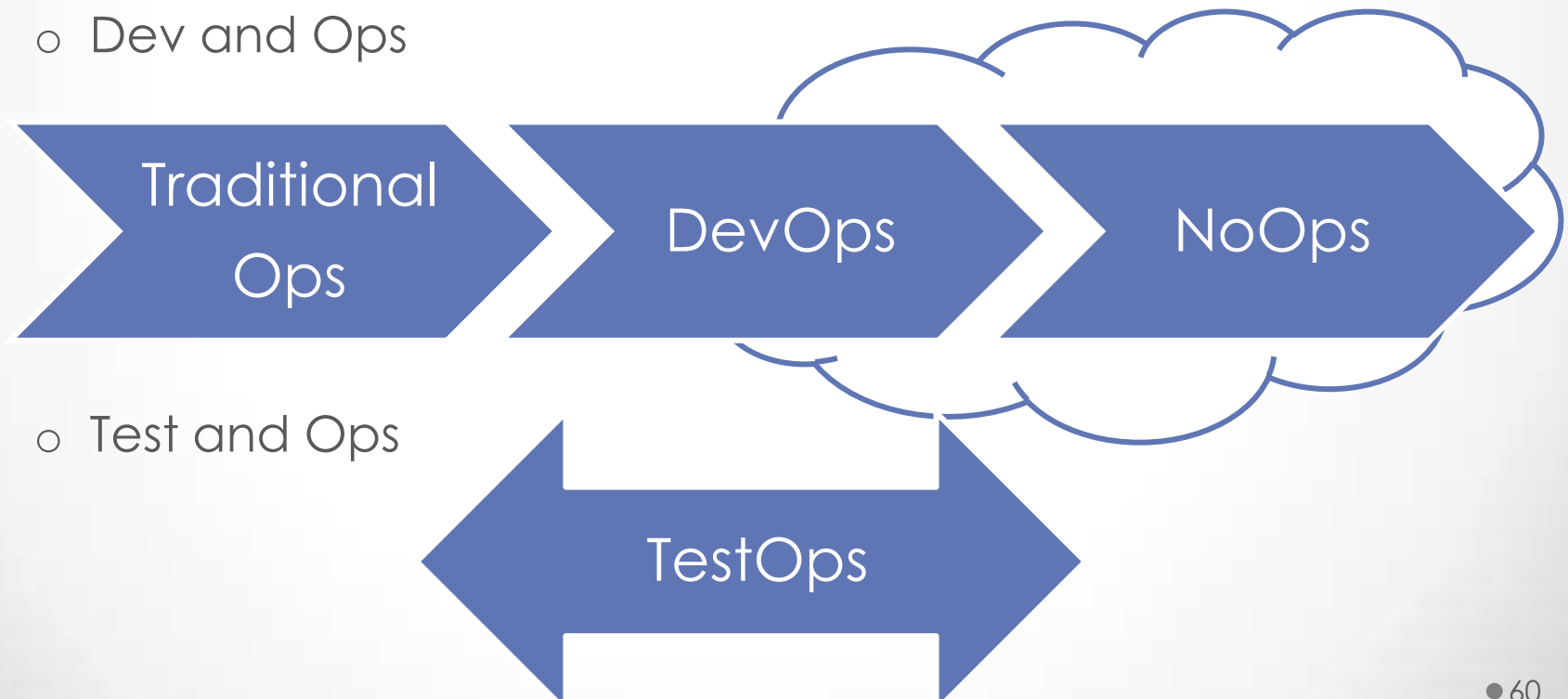
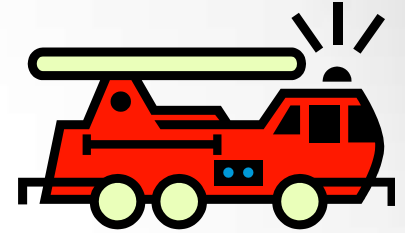
- 30 Engineers on a con-call

To

- 1.5 engineers and automation

Operations

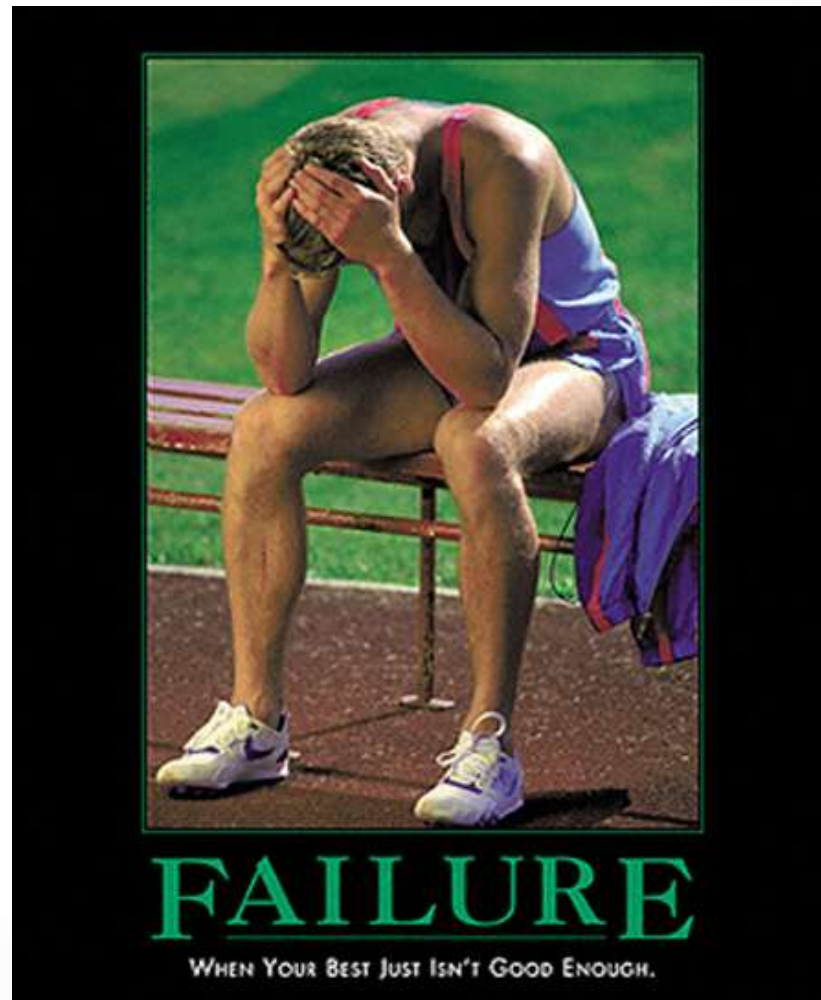
- Ops runs the environment
- Ops needs to be in the loop on TiP
 - Else they may react as if a real problem were occurring
- Need to define roles and responsibilities
 - Dev and Ops



- Test and Ops

Destructive Testing in Production

- Google first year of a new data center [Google DC, 2008]
 - 20 rack failures, 1000 server failures and thousands of hard drive failures
- High Availability means you must Embrace Failure
 - How do you test this?



Netflix Tests its “Rambo Architecture”

- ...system has to be able to succeed, no matter what, even all on its own
- Test with **Fault Injection**



[Netflix Army, July 2011]

- Netflix *Simian Army*
 - **Chaos monkey** randomly **kills** production instance in AWS
 - **Chaos Gorilla** simulates an outage of an entire Amazon AZ
 - Janitor Monkey, Security Monkey, Latency Monkey.....





April 2011



TiP A-Z

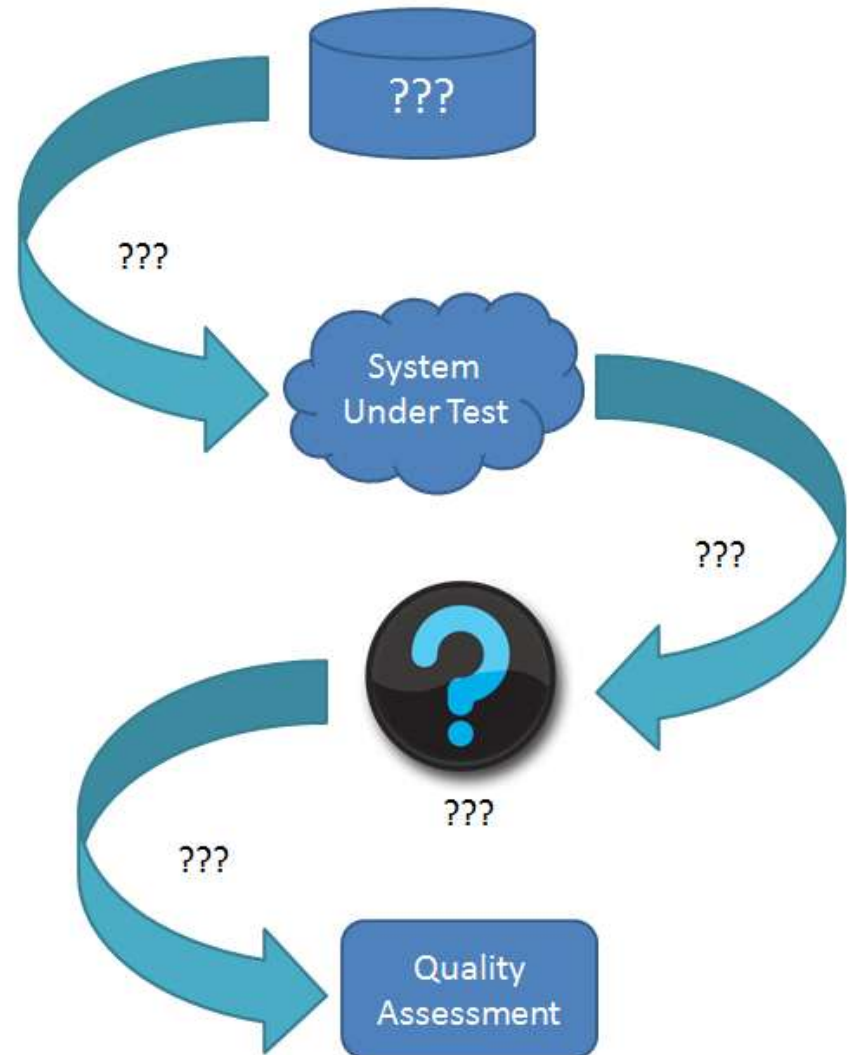
A	Active Monitoring	N	Not to Do, What
B	Big Data	O	Operations
C	Client Instrumentation	P	Passive Monitoring
D	Data Mining	Q	Quality Signal
E	Experimentation	R	Real Data Input
F	Fault Injection	S	System Stress
G	Go / No-go for deployment	T	TestOps
H	High Availability	U	Users
I	Iterative Virtuous Cycle	V	3 V's of Big Data
J	JSI: J-script Instrumentation	W	Write Once, Test Anywhere
K	Killing production instances	X	eXposure Control
L	Load Testing in Production	Y	Y TiP?
M	Most Interesting Man in the World	Z	Zymurgy

Changing the Quality Signal

...

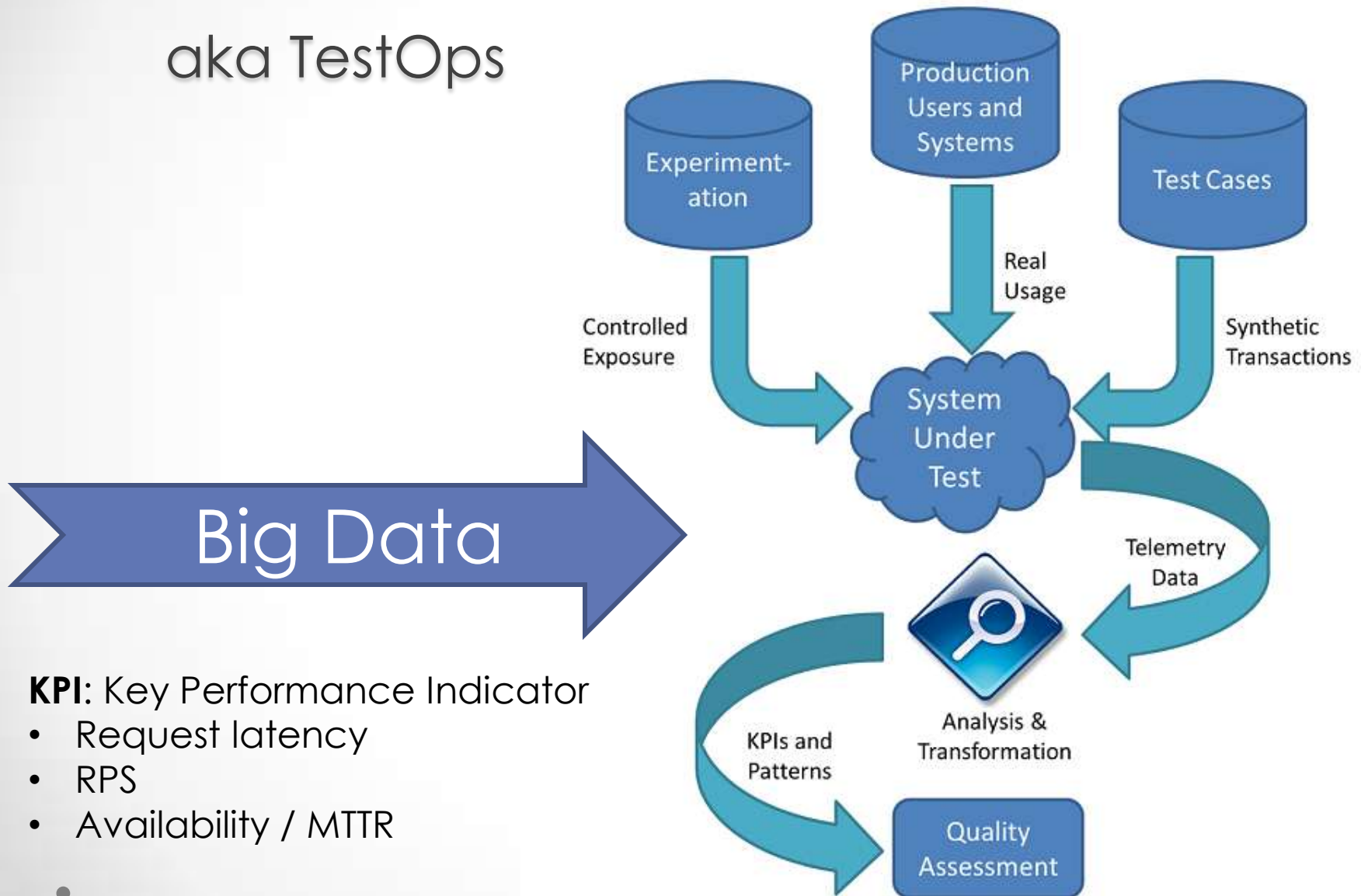
What Can Testing Be?

Change the signal from
Test Results to...



Big Data Quality Signal

aka TestOps



KPI: Key Performance Indicator

- Request latency
- RPS
- Availability / MTTR

TiP A-Z

A	Active Monitoring	N	Not to Do, What
B	Big Data	O	Operations
C	Client Instrumentation	P	Passive Monitoring
D	Data Mining	Q	Quality Signal
E	Experimentation	R	Real Data Input
F	Fault Injection	S	System Stress
G	Go / No-go for deployment	T	TestOps
H	High Availability	U	Users
I	Iterative Virtuous Cycle	V	3 V's of Big Data
J	JSI: J-script Instrumentation	W	Write Once, Test Anywhere
K	Killing production instances	X	eXposure Control
L	Load Testing in Production	Y	Y TiP?
M	Most Interesting Man in the World	Z	Zymurgy

What NOT to Wear!



What Not to Do

...

Risks of Synthetic Data...

Who's Got the Monkey?

Who's Got The Monkey Now?

....and Mitigations

we've setup a test order system specifically to handle these requests, and no testing orders made it into the sales data



TiP is a common and costly technical malpractice

- Because both [prod and test]... ran on the same base WebSphere Application Server
 - their logs are shared
 - any upgrade to the SDK would disrupt both application servers

“Simply, you should have a separate test system that is identical to your production system”



[IBM, 2011]

What Not to Do, IBM

- Disrupt user experience introduced by testing: for example: outages
- Failure to understand the production environment and the effect of TiP on it.
- Co-mingling/corrupting production data



Wireless Mouse?

amazon.com **Prime**

Hello, Seth Eliot. We have [recommendations](#) for you. (Not Seth?)

Seth's Amazon.com | [Today's Deals](#) | [Gifts & Wish Lists](#) | [Gift Cards](#)

The All-New Kindle: Only \$139

Your Account | Help

Shop All Departments ▾ Search Electronics ▾ GO [Cart](#) [Your Lists](#) ▾

[All Electronics](#) | [Brands](#) | [Bestsellers](#) | [Audio & Home Theater](#) | [Camera & Photo](#) | [Car Electronics & GPS](#) | [Cell Phones & Accessories](#) | [Computers](#) | [MP3 Players](#) | [TV & Video](#) | [Deals](#)

Prime

Member: Seth Eliot

One-Day 1-Click®—\$3.99 Two-Day 1-Click®—FREE

Ship to: Seth Eliot ▾

Note: This item is not immediately available to ship.

Electronics Test ASIN - 207

by [test pubcode](#)

★★★★☆ (3 customer reviews) | [Like](#) (6,592,029)

Price **\$99,999.00** & eligible for free shipping with **Amazon Prime**

Temporarily out of stock.

Order now and we'll deliver when available. We'll e-mail you with an estimated delivery date as soon as we have more information. Your account will only be charged when we ship the item. Ships from and sold by **Amazon.com**. Gift-wrap available.



Customer image from [Jason Dickinson "Jason"](#)

Quantity: 1 ▾

[Add to Cart](#)

or Buy now with

[Two-Day 1-Click®—FREE](#)

Ship to: Seth Eliot ▾

☐ Add gift-wrap/note

or

[Buy with DayPhrase](#)

[Add to Wish List](#) ▾

[Add to Shopping List](#)

[Add to Baby Registry](#)

What Not to Do, Amazon

- Exposure of test artifacts to end users: e.g. exposed test items, test emails sent
- Misuse of PII customer data
- Leaking sensitive new features prior to official launch

Amazon's Digital Video sneak peek

Amazon's Digital Video sneak peek: "Amazon Unbox"

It seems Amazon might soon be launching their digital video download store, called **"Amazon Unbox Video"**. According to what I can find, it'll have purchase and rental capability, as well as support for devices other than your PC (Your TV and Creative Zen Vision at least). They also have a standalone video player, somewhat like iTunes (Windows XP and Windows NT only it seems)

Here are screenshots of the pages and the player. Click pic to see it large.

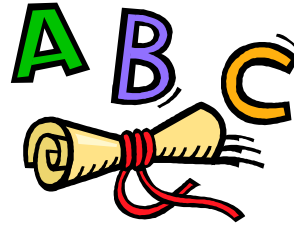
Main page:



[Kokogiak, 2006]

TiP A-Z

A	Active Monitoring	N	Not to Do, What
B	Big Data	O	Operations
C	Client Instrumentation	P	Passive Monitoring
D	Data Mining	Q	Quality Signal
E	Experimentation	R	Real Data Input
F	Fault Injection	S	System Stress
G	Go / No-go for deployment	T	TestOps
H	High Availability	U	Users
I	Iterative Virtuous Cycle	V	3 V's of Big Data
J	JSI: J-script Instrumentation	W	Write Once, Test Anywhere
K	Killing production instances	X	eXposure Control
L	Load Testing in Production	Y	Y TiP?
M	Most Interesting Man in the World	Z	Zymurgy



Summary

...

The latest version of this slide deck can be found at:

http://bit.ly/seth_stp_oct2012

TiP A-Z

A	Active Monitoring	N	Not to Do, What
B	Big Data	O	Operations
C	Client Instrumentation	P	Passive Monitoring
D	Data Mining	Q	Quality Signal
E	Experimentation	R	Real Data Input
F	Fault Injection	S	System Stress
G	Go / No-go for deployment	T	TestOps
H	High Availability	U	Users
I	Iterative Virtuous Cycle	V	3 V's of Big Data
J	JSI: J-script Instrumentation	W	Write Once, Test Anywhere
K	Killing production instances	X	eXposure Control
L	Load Testing in Production	Y	Y TiP?
M	Most Interesting Man in the World	Z	Zymurgy

Time for a Beer...

Zymurgy

zy · mur · gy (z mûr j) n. The branch of chemistry that deals with fermentation processes, as in brewing.



References

[Google Talk, June 2007]	Google: Seattle Conference on Scalability: Lessons In Building Scalable Systems, Reza Behforooz http://video.google.com/videoplay?docid=6202268628085731280
[Unpingco, Feb 2011]	Edward Unpingco; Bug Miner; Internal Microsoft Presentation, Bing Quality Day
[Barranco, Dec 2011]	René Barranco; Heuristics-Based Testing; Internal Microsoft Presentation
[Dell, 2012]	http://whichtestwon.com/dell%e2%80%99s-site-wide-search-box-test
[Microsoft.com, TechNet]	http://technet.microsoft.com/en-us/library/cc627315.aspx
[Cockcroft, March 2012]	http://perfcap.blogspot.com/2012/03/ops-devops-and-noops-at-netflix.html
[Deschamps, Johnston, Jan 2012]	Experiences of Test Automation; Dorothy Graham; Jan 2012; ISBN 0321754069; Chapter: "Moving to the Cloud: The Evolution of TiP, Continuous Regression Testing in Production"; Ken Johnston, Felix Deschamps
[Google DC, 2008]	http://content.dell.com/us/en/gen/d/large-business/google-data-center.aspx?dgc=SM&cid=57468&lid=1491495 http://perspectives.mvdirona.com/2008/06/11/JeffDeanOnGoogleInfrastructure.aspx
[Kohavi, Oct 2010]	Tracking Users' Clicks and Submits: Tradeoffs between User Experience and Data Loss http://www.exp-platform.com/Pages/TrackingClicksSubmits.aspx
[Strata Jan 2012]	What is big data? - An introduction to the big data landscape http://radar.oreilly.com/2012/01/what-is-big-data.html

References, continued

[Netflix Army, July 2011]	The Netflix Simian Army; July 2011 http://techblog.netflix.com/2011/07/netflix-simian-army.html
[Google-Wide Profiling, 2010]	Ren, Gang, et al. Google-wide Profiling: A Continuous Profiling Infrastructure for Data Centers. [Online] July 30, 2010. research.google.com/pubs/archive/36575.pdf
[Facebook ships, 2011]	http://framethink.blogspot.com/2011/01/how-facebook-ships-code.html
[Google BusinessWeek, April 2008]	How Google Fuels Its Idea Factory, BusinessWeek, April 29, 2008; http://www.businessweek.com/magazine/content/08_19/b4083054277984.htm
[IBM 2011]	http://www.ibm.com/developerworks/websphere/techjournal/1102_supauth/1102_supauth.html
[Kokogiak, 2006]	http://www.kokogiak.com/gedankengang/2006/08/amazons-digital-video-sneak-peek.html
[Google GTAC 2010]	Whittaker, James. GTAC 2010: Turning Quality on its Head. [Online] October 29, 2010. http://www.youtube.com/watch?v=cqwXUTjcabs&feature=BF&list=PL1242F05D3EA83AB1&index=16 .
[Google, JW 2009]	http://googletesting.blogspot.com/2009/07/plague-of-homelessness.html
[STPCon, 2012]	STPCon Spring 2012 - Testing Wanted: Dead or Alive – March 26, 2012
[Cook, June 2010]	Ganglia, OSD: Cook, Tom. A Day in the Life of Facebook Operations. Velocity 2010. [Online] June 2010. http://www.youtube.com/watch?v=T-Xr_PJdNmQ

Contact Information

Seth Eliot

Senior Knowledge Engineer in Test
Microsoft

seth.eliot@microsoft.com

Twitter: setheliot@

Blog: http://bit.ly/seth_qa

Thanks for Attending

Please fill out an evaluation form and drop it in the collection basket located at the back of the room.

Session 103

A to Z Testing in Production

Seth Eliot