

Debugging @ Scale

Kumar Srinivasamurthy
Bing & Cortana Engineering



conduct 2.130676415
Relays 6-2 in 033 failed special speed test
in relay " 10,000 test.

Relay
2145
Relay 3371

1100 Started Cosine Tape (Sine check)
1525 Started Mult + Adder Test.

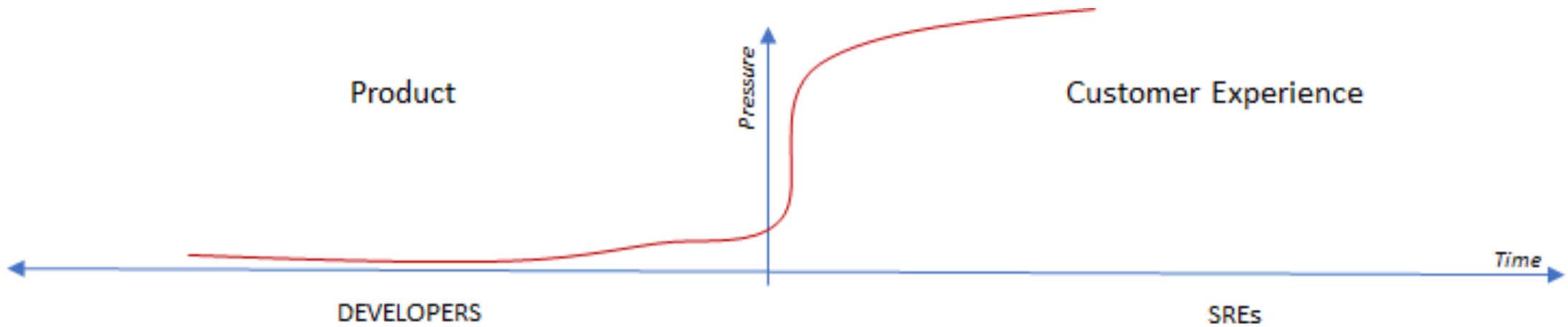
1545



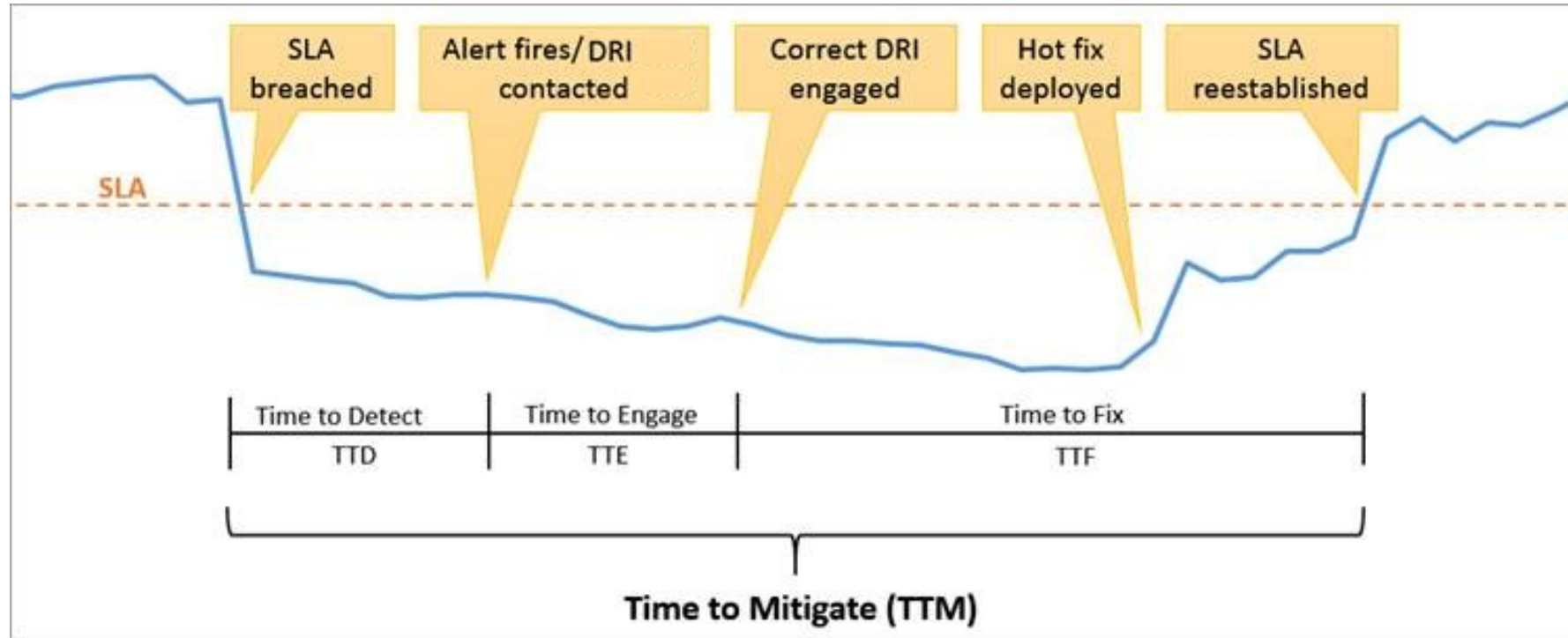
Relay #70 Panel F
(moth) in relay.

~~1630~~ 1630 Antangent started.
1700 closed down.
First actual case of bug being found.

Looks familiar?



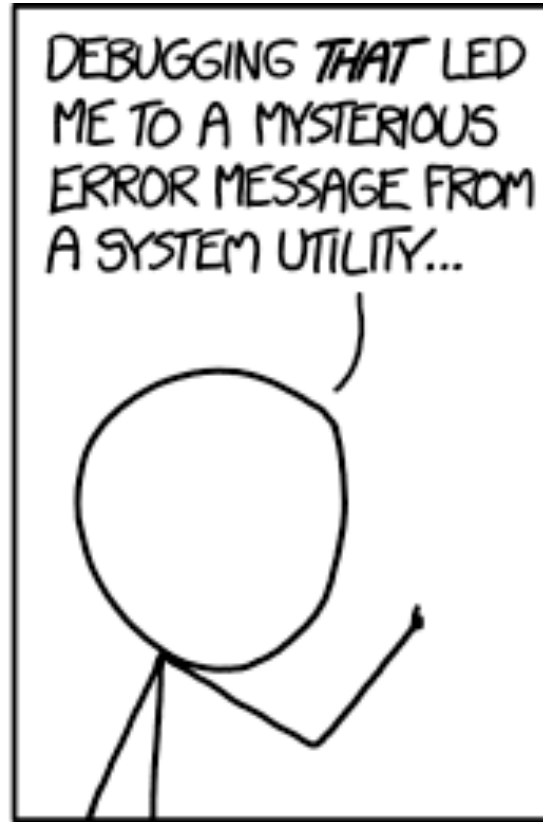
The (dual) nature of debugging



When auto-mitigation fails...



TURNS OUT IT WASN'T THE BROWSER—THE ISSUE WAS WITH MY KEYBOARD DRIVER.



<https://xkcd.com/1722>

Debugging is Hard



Who most of us at ❤️ are - Explorers



Who we should Be? - ER

Expert Debugger

```

D:\Apache\htdocs\site\checklogin.php
3 $allow = false;
4 $lines = file('ip_allow.lst');
5 $lines[count($lines)] = '127.0.0.*';
6
7 for ($i=0; $i<count($lines); $i++) {
8     if (ereg(ereg_replace('\././\./', '\.', e
9         $allow = true;
10         break;
11     }
12 }
13
14 if ($allow == false) {
15     die('allow = true <ss deny.</h1>');
16 }
17
18 include_once('session.php');
19 if ($SESSION['login'] == '') {
20     $SESSION['REFERER'] = $SERVER['RE
21     header("Location: http://".$_SERVER
22 }
23
24

```

Variables

Name	Value
\$allow	True
Group: \$_REQUEST	
Group: \$argv	
Group: \$_HTTP_COOKIE_VARS	
Group: \$_HTTP_GET_VARS	
Group: \$_HTTP_SERVER_VARS	
\$_HTTP_SERVER_CGI/1.1	
\$_HTTP_SERVER_localhost	

Output

1. Source	2. Browser	3. Profile	Hit co...	Time ...
checklogin.php			3	-0.04
checklogin.php			4	-1.49
checklogin.php			5	-0.02

DEBUGGER

JT Demo-RayTracer.exe - Telerik JustTrace

Start Session

Back Forward Get Snapshots Compare Snapshots Show Timeline Demo-RayTrac Kill & Get Kill Profile Force GC

NAVIGATION 13:17:17 13:17:22 13:17:27 13:17:32 13:17:37 13:17:42 13:17:47

5000kb
4000kb
3000kb
2000kb
1000kb
0kb

Current views: Snapshot at 13:17:41

Overview

- Metrics
- Snapshots Summary
- Assemblies
- Potential Binding Leaks In...
- Disposed Objects
- Event-Handler Analysis
- Instances
- By Type
- Largest Memory Retainers
- Allocation Call Trees
- All Threads

Memory Profiling Session Using the Timeline

CPU USAGE GRAPH

MEMORY USAGE GRAPH (KB)

SESSION SUMMARY

Profiler Type: Memory
Started: 08/27/13 01:17

PERFORMANCE COUNTERS

- CPU Usage: 48.87%
- Working Set: 34.9 MB
- Memory Usage: 367 bytes
- Gen 0 GC: 8
- Gen 2 GC: 3
- Included GC: 2
- Time in GC: 0.28%
- Time in JIT: 0.00%
- Contentions: 0
- Contentions/sec: 0
- Exceptions: 0
- Exceptions/sec: 0
- Managed Heap Total Size: 4,35 KB

Some memory counters are updated only after GC.

MEMORY PROFILER

All Systems - Htdp login:cs

Dashboard Events Account Help -> Mx

Oct 02 13:00:14 [INFO] [Oct 02 14:00:14] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:14] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

Oct 02 13:00:14 [INFO] [Oct 02 14:00:14] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:14] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:15 [INFO] [Oct 02 14:00:15] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:15] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

Oct 02 13:00:15 [INFO] [Oct 02 14:00:15] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:15] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:16 [INFO] [Oct 02 14:00:16] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:16] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:16 [INFO] [Oct 02 14:00:16] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:16] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

Oct 02 13:00:17 [INFO] [Oct 02 14:00:17] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:17] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:17 [INFO] [Oct 02 14:00:17] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:17] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

Oct 02 13:00:18 [INFO] [Oct 02 14:00:18] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:18] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:18 [INFO] [Oct 02 14:00:18] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:18] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

Oct 02 13:00:19 [INFO] [Oct 02 14:00:19] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:19] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:19 [INFO] [Oct 02 14:00:19] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:19] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

Oct 02 13:00:20 [INFO] [Oct 02 14:00:20] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:20] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:20 [INFO] [Oct 02 14:00:20] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:20] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

Oct 02 13:00:21 [INFO] [Oct 02 14:00:21] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:21] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:21 [INFO] [Oct 02 14:00:21] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:21] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

Oct 02 13:00:22 [INFO] [Oct 02 14:00:22] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:22] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:22 [INFO] [Oct 02 14:00:22] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:22] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

Oct 02 13:00:23 [INFO] [Oct 02 14:00:23] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:23] [4000] POST /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 302 ...

Oct 02 13:00:23 [INFO] [Oct 02 14:00:23] [logon] [a] [101.101.98.99] ... [102/01/2014:14:00:23] [4000] GET /wp/press-there-unavailable-12/factor=invalid_email HTTP/1.1 200 2316 ...

LOGS

Traditional Tools of the Debugging Trade





Traditional tools **don't scale** 😊

Concepts we can use...



MAKE ERROR MESSAGES
USEFUL FOR USERS *AND*
ENGINEERS



500

Unexpected Error :(

An error occurred and your request couldn't be completed. Please try again.

A common error message



500 Internal Server Error

Sorry, something went wrong.

A team of highly trained monkeys has been dispatched to deal with this situation.

If you see them, send them this information as text (screenshots frighten them):

```
APCGX7pzux_z-HXU9ctCRbdKy09R0ID3xzIqQrE9RFG5usbJjN_fyaUG
9RRFTEE3D71REBsmPjSI4vuPMEQcrOdXE2gUchPjBwamxjfuQ5nvmshj
1iowFfwAYA2tasi2oHy3DbNDWQFQTvx8JnSeMM1nhU9bIYL_RNLqISp5
be2peCtgeKNr3mBtVD2pJQqaqfCkv47f1I29E-7Ujun57VHEVEFW4CYL
7019s_Nuyo4pCbX7YThW-vTEiv4QLrKTXuL_V8wDbo7aJ88og8zBN9h0
mxSUSEweC3uAENVz-pEu6BvKN12VtBpE4Ru2Z0RebkZ2ETuR_2g6XS6s
cHFBQOgXjU-B4uiDeR_GzWYhqGuVbHLm2ef1Abqau2oQqD1bVYoIdj-2
7FrLaT_soe6MV95pfDpapSed5LW4iiZRM4BGS-9HEeHvqV10gm-1sxtq
_0RJL6Xte0Ke-IdxR97ddtTiQwEVvSFtE1faVfb07331zXZPri6BLVPK
pXe4uSSgseYwU0Dk20fXiCB8KwbPajcpdnjMGA9GPOkSt_lvAFF5UyCr
oU5_FnmHv145Q-6oHwybw90hbRnb4ECj3smIuN6J8vO4tqLG0HtI00x7
s9AaFYDAYnH6EPd_1ZKIgg_uXWn61-Sj_EaAF8vGAhPwZIFSZmMyAITe
fbFD0v1Lqps-OjdqX-yP0nfPcw9pHKY-WHF47azrIoL6MnJ3IyJ1ML_0
wayWt4rr2hjrUTD1vYyyiULF9L1xoffLemw8pNqasSvJ4hLW9ReBws9
```



It's not you, it's us

Bing isn't available right now, but everything should be back to normal very soon.

Ref A: 70F2DAA96DF64F87A2F372FC4748F3D2 Ref B: WSTEDGE0519 Ref C: Mon May 1 15:37:36 2017 PST



A better approach



USE DISTRIBUTED
TRACING AS THE NEW
'DEBUGGER'

STAMP Manager

1E4637D6F6C4ED4A78EE1818175469

Create XAP Session | Edge Logs | SNR SPeef Logs Viewer | Traffic Dashboard

IndexGenefanking1-Prod-CH01 Total: 65	AnswersMultimedia-Prod-CH1b Total: 9	AnswersNews-Prod-CH1b Total: 1	ContUX-Prod-CH1b Total: 17
Scalix-ContUX-Prod-CH1b Total: 82	XAP-Prod-CH1b Total: 2	AnswersReference-Prod-CH1b Total: 16	ObjectStoreMulti-Prod-CH1b Total: 140
ObjectStoreMulti-Prod-Cont Total: 3	ScalSocial-Prod-CH1 Total: 2	Cache-Prod-CH1b Total: 6	Scalix/MM-Prod-CH1b Total: 4
IndexGenefanking1-Prod-CH1b Total: 174	Scalix/20WFrash-Prod-CH1b Total: 25	Kusto-IntentsHub-Prod-CH1 Total: 651	Scalix/News-Prod-CH1 Total: 2
MMGenes2-Prod-CH1b Total: 4	IndexGenesSelection-Prod-CH1b Total: 3	IndexGenesSelection1-Prod-CH01 Total: 3	Kusto-SpeechEAP-Prod-CH1b Total: 1

Total: 1219 Loaded: 80

Find on page:

Log Levels: Error Warn Info Debug Status Info

IndexGenefanking1-Prod-CH01 TC 0:014:51970:334	T:\A\AnsoPerf_138262.log
IndexGenefanking1-Prod-CH01 TC 0:014:51970:334	T:\A\AnsoPerf_135158.log
IndexGenefanking1-Prod-CH01 TC 0:014:51970:383	T:\A\AnsoPerf_164969.log
IndexGenefanking1-Prod-CH01 TC 0:014:51970:382	T:\A\AnsoPerf_171904.log
IndexGenefanking1-Prod-CH01 TC 0:014:51970:382	T:\A\AnsoPerf_114141.log
IndexGenefanking1-Prod-CH01 TC 0:014:51970:348	T:\A\AnsoPerf_120887.log
IndexGenefanking1-Prod-CH01 TC 0:014:51970:311	T:\A\AnsoPerf_114884.log
IndexGenefanking1-Prod-CH01 TC	T:\A\AnsoPerf_131785.log



Zipkin Investigate system behavior Find a trace Realtime Aggregates

Duration: 195.000ms Services: 8 Depth: 2 Total Spans: 17

Expand All Collapse All Filter Service Search

cms-lookup-server x1 oppdragsserver x1 prod x1 search-management x3 smajobber-thrift-server x1 smajobberlookupservice x1 statistics x1 sybase-finnst x9

Services	39.000ms	78.000ms	117.000ms	156.
prod	195.000ms : /finn/			
cms-lookup-server	15.000ms : CmsLookupService			
sybase-finnst	5.000ms : pr_myads_pagination_ordered			
statistics	2.000ms : AdCountService			
oppdragsserver	6.000ms : OppdragService			
smajobberlookupservice	6.000ms : SmajobberLookupService			
search-management	2.000ms : SearchManagementService			
search-management	3.000ms : SearchManagementService			
search-management	9.000ms : SearchManagementService			
sybase-finnst	2.000ms : pr_person_fetch			
sybase-finnst	2.000ms : pr_org_fetchauxaddress			
sybase-finnst	2.000ms : pr_user_preferences_fetch			
sybase-finnst	1.000ms : pr_user_profile_fetch			
sybase-finnst	7.000ms : pr_person_update			
sybase-finnst	2.000ms : pr_user_preferences_ins_upd			
sybase-finnst	2.000ms : pr_user_profile_ins_upd			
sybase-finnst	2.000ms : pr_org_insertauxaddress			

APPLY TRANSFORMS AND
'MACHINE LEARNING' TO
MAKE LOGS USEFUL

THIS IS YOUR MACHINE LEARNING SYSTEM?

YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.





...From Our 'Pile Of Linear Algebra'

- Log-relevant Tokenization
- Feature Extraction
- Instance-Based Learning
 - KNN: k-Nearest Neighbor
- Linear Classifiers
- Clustering
 - k-Means



Design for Reproducibility

Reproducibility

- Log playback
- Forking traffic
- Single box/Mocking
- Snapshot Debugger
- Experimentation



INVEST IN DEEP OPTICS WITH NRT METRICS

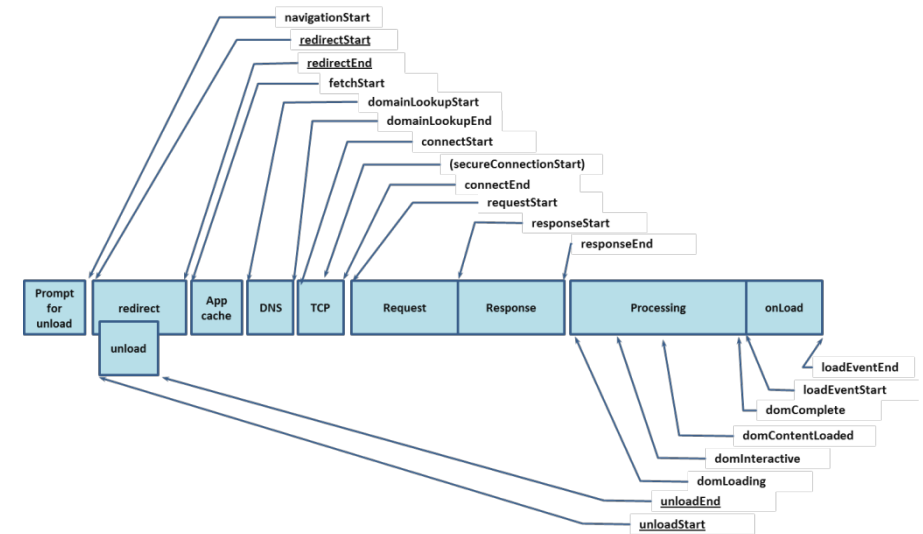


NEAR REAL TIME METRICS



➤ Expand to business metrics

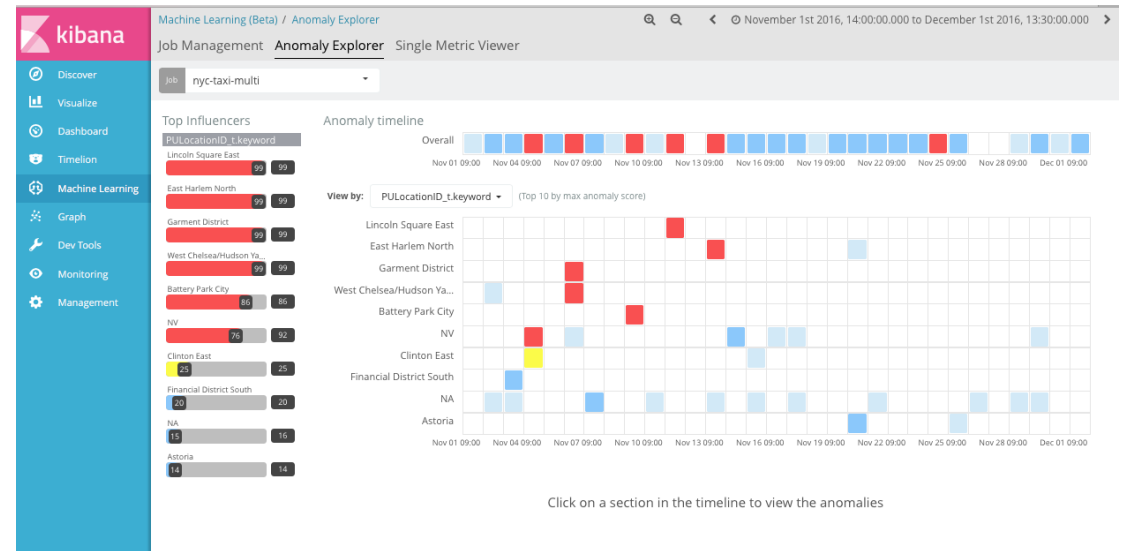
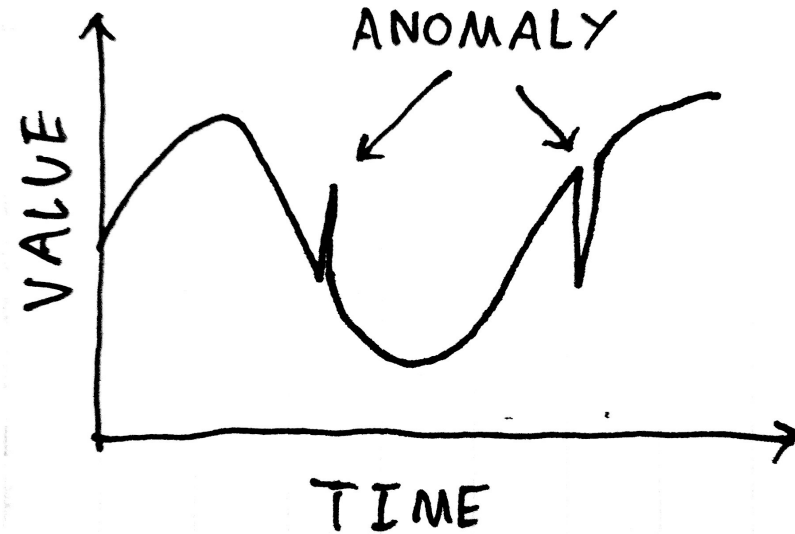
REAL USER MEASUREMENTS (RUM)



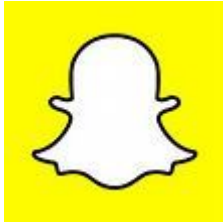
Navigation Timing API

DEVELOP ANOMALY
DETECTION AS A
PLATFORM CAPABILITY

- Detect anomalies across metrics
- Enable outlier detection on logs
- High data dimensionality allows better detection



PRIVACY *HAS* TO BE A
FIRST-CLASS CITIZEN FOR
DEBUGGING



Be Mindful Of:

- GDPR – Very pertinent for EU
- Log sanitization
- Log Retention and On demand purge
- Exposure to 3rd Party Analysis Systems and Hosting
- The Latent Power of Aggregated Logs



OLD TRICKS STILL WORK,
BUT CAN BE TRICKY

Still in use...

- Debug Mode for subset in Production
- Deploying private builds on a subset
- Logging onto machine instances
- Attaching debugger
- Others?



CLOSING THOUGHTS

- NRT Telemetry and Distributed Tracing – Key Elements
- At scale, machine-learning isn't just a buzzword – it *actually* delivers
- Be mindful of privacy concerns. Your business is at stake
- Debugging in fun 😊

Only half of programming is coding. The other 90% is debugging.

/r/ProgrammerHumor

Questions?

- kumars@microsoft.com
- [@00kumars](#)

