

Scaling with Four Million Users

Simon Engledew

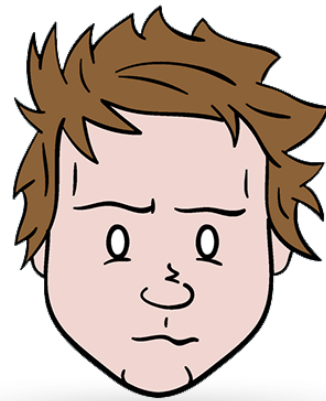


MyMaths Platform Tech Team

(we are tiny)



Simon Engledew
Senior Engineer



Darren Royle
Platform Manager



Ed Smith
Engineer

Product rewrite



Getting ready for production

6 months to go, heads on the block...

The image displays two overlapping screenshots of the MyMaths.co.uk web application, illustrating the user interface for managing classes and tasks.

Left Screenshot: All Classes

The URL is `localhost:3000/assessment/allocation/index`. The page shows a table of classes with columns for Class, Students, and Tasks. A circular callout highlights the class selection options:

- Class
- Secondary class
- Secondary class 2

Right Screenshot: Secondary school settings

The URL is `localhost:3000/assessment/secondary-school/settings`. The page shows a table of tasks with columns for Topic, Task name, and secondary s. The table is filtered by 'Secondary class'.

Topic	Task name	secondary s
1	Number	Counting On and Back
1	Number	Number Facts and Doubles...
2	Number	Number Bonds
2	Number	Number Facts and Doubles...
3	Number	Number Facts and Doubles...
4	Number	Number Facts and Doubles...

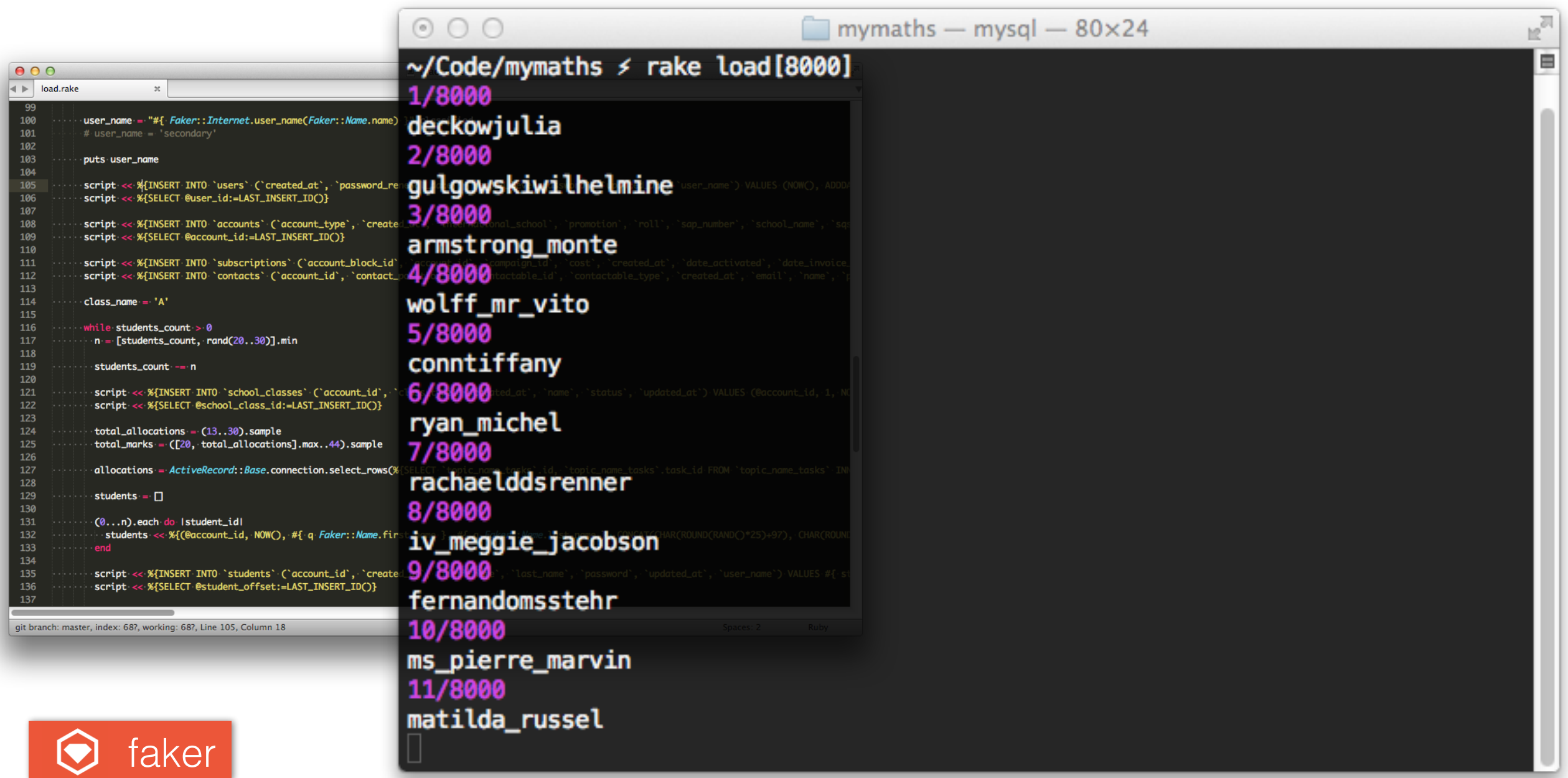
The sidebar on the right allows filtering tasks by:

- Selected Dates
- Level
- Type of Task
- Library
- Boosters
- GCSE Statistics

Buttons for 'Download results', 'Delete these results', 'Message students', and 'Toggle Fullscreen' are visible above the table.

Generate Realistic Test Data

(300 million records should do)



The image displays two terminal windows. The left window, titled 'load.rake', shows a Ruby script using the Faker gem to generate test data. The script includes comments and code for inserting records into 'users', 'accounts', 'subscriptions', 'contacts', 'school_classes', and 'students' tables. The right window, titled 'mymaths — mysql — 80x24', shows the output of the 'rake load[8000]' command, displaying a list of generated names and their corresponding IDs (e.g., 1/8000 deckowjulia, 2/8000 gulgowskiwilhelmine, etc.).

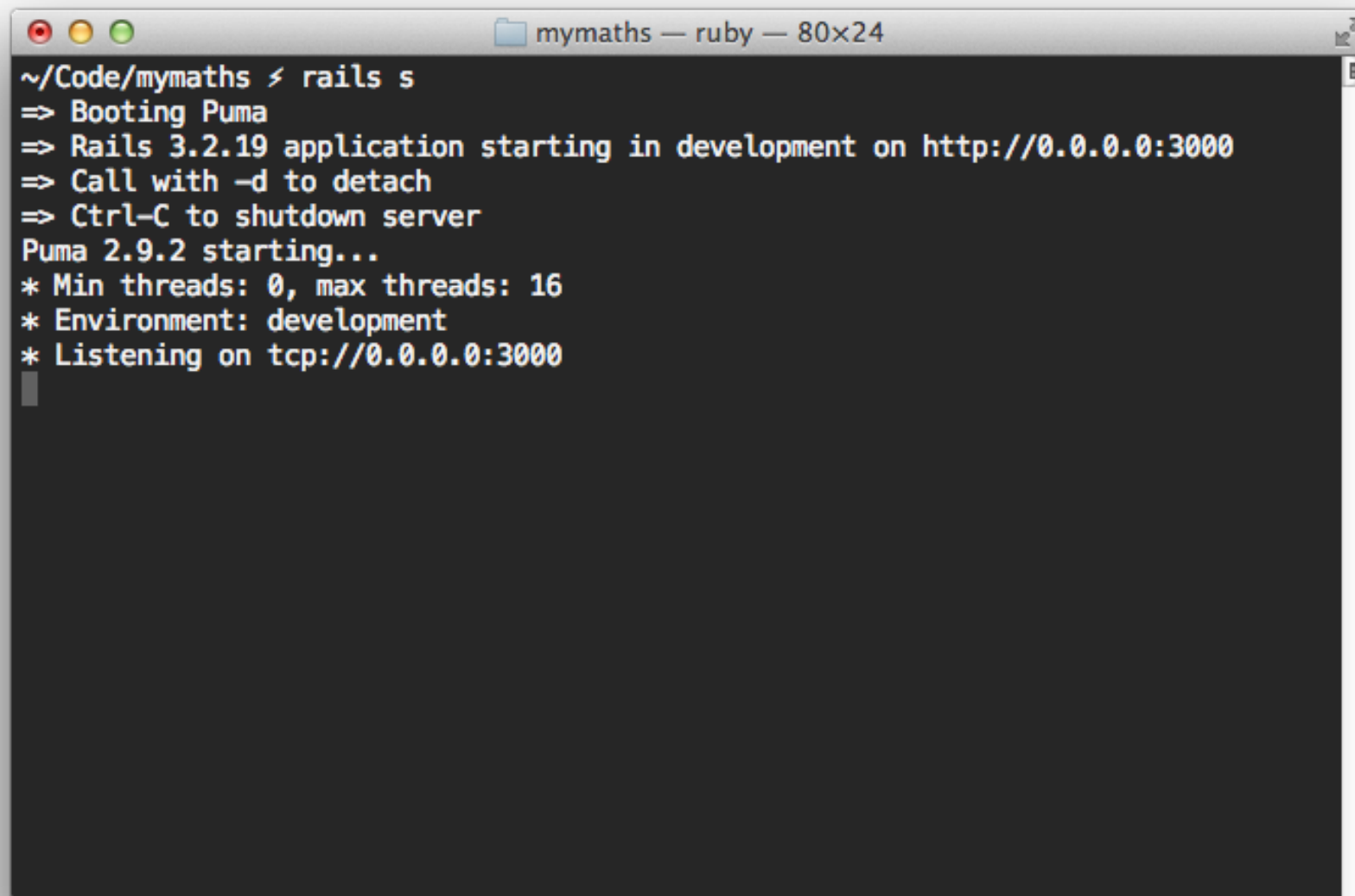
```
load.rake
99
100 user_name = "#{ Faker::Internet.user_name(Faker::Name.name) }"
101 # user_name = 'secondary'
102
103 puts user_name
104
105 script << %{\INSERT INTO `users` (`created_at`, `password_resets`) VALUES (NOW(), ADDO
106 script << %{\SELECT @user_id:=LAST_INSERT_ID()}
107
108 script << %{\INSERT INTO `accounts` (`account_type`, `created_at`, `initial_school`, `promotion`, `roll`, `sap_number`, `school_name`, `sq
109 script << %{\SELECT @account_id:=LAST_INSERT_ID()}
110
111 script << %{\INSERT INTO `subscriptions` (`account_block_id`, `campaign_id`, `cost`, `created_at`, `date_activated`, `date_invoice
112 script << %{\INSERT INTO `contacts` (`account_id`, `contactable_id`, `contactable_type`, `created_at`, `email`, `name`, `p
113
114 class_name = 'A'
115
116 while students_count > 0
117   n = [students_count, rand(20..30)].min
118
119   students_count -= n
120
121   script << %{\INSERT INTO `school_classes` (`account_id`, `school_class_id`, `name`, `status`, `updated_at`) VALUES (@account_id, 1, N
122   script << %{\SELECT @school_class_id:=LAST_INSERT_ID()}
123
124   total_allocations = (13..30).sample
125   total_marks = ([20, total_allocations].max..44).sample
126
127   allocations = ActiveRecord::Base.connection.select_rows(%{\SELECT `topic_name_tasks`.`id`, `topic_name_tasks`.`task_id` FROM `topic_name_tasks` IN
128
129   students = []
130
131   (0..n).each do |student_id|
132     students << %{\(@account_id, NOW(), #{ q Faker::Name.first_name }, #{ q Faker::Name.last_name }, #{ q Faker::Name.password }, CHAR(ROUND(RAND()*25)+97), CHAR(ROUND
133   end
134
135   script << %{\INSERT INTO `students` (`account_id`, `created_at`, `last_name`, `password`, `updated_at`, `user_name`) VALUES #{ st
136   script << %{\SELECT @student_offset:=LAST_INSERT_ID()}
137
```

```
mymaths — mysql — 80x24
~/Code/mymaths ✂ rake load[8000]
1/8000 deckowjulia
2/8000 gulgowskiwilhelmine
3/8000 armstrong_monte
4/8000 wolff_mr_vito
5/8000 conntiffany
6/8000 ryan_michel
7/8000 rachaelddsrenner
8/8000 iv_meggie_jacobson
9/8000 fernandomsstehr
10/8000 ms_pierre_marvin
11/8000 matilda_russel

```



Started up the site to have a look...

A terminal window titled "mymaths — ruby — 80x24" with standard macOS window controls (red, yellow, green buttons). The terminal shows the command "rails s" being executed in the directory "~/Code/mymaths". The output includes instructions for booting Puma, the Rails version (3.2.19), the development environment, and the server listening on http://0.0.0.0:3000. Puma 2.9.2 is shown starting with 16 max threads.

```
~/Code/mymaths ✗ rails s
=> Booting Puma
=> Rails 3.2.19 application starting in development on http://0.0.0.0:3000
=> Call with -d to detach
=> Ctrl-C to shutdown server
Puma 2.9.2 starting...
* Min threads: 0, max threads: 16
* Environment: development
* Listening on tcp://0.0.0.0:3000
█
```

Boom!

(...don't panic)

MyMaths - Bringing maths alive

Help Log out Search...

Allocation Results Admin

OXFORD

All Classes

View Results Download Markbook

Loading...

Action Controller: Exception

localhost:3000/assessment/class_admin/1028758195

Simon

ActiveRecord::StatementInvalid in Assessment::AdminController#class_admin_index

Timeout::Error: execution expired: SELECT students.id AS student_id, students.first_name AS forname, students.last_name AS lastName, students.user_name AS loginNumber, CONCAT_WS(, students.first_name, students.last_name) AS name, GROUP_CONCAT(outer_school_classes.name SEPARATOR ', ') AS class, school_classes.id AS school_class_id, students.password, students.id AS edit FROM `students` INNER JOIN `student_classes` ON `student_classes`.`student_id` = `students`.`id` INNER JOIN `school_classes` ON `school_classes`.`id` = `student_classes`.`school_class_id` INNER JOIN student_classes AS outer_student_classes ON outer_student_classes.student_id = students.id INNER JOIN school_classes AS outer_school_classes ON outer_school_classes.id = outer_student_classes.school_class_id WHERE `students`.`account_id` = 1030095999 AND `school_classes`.`id` = 1028758195 GROUP BY students.id ORDER BY last_name, first_name

Rails.root: /Users/engledes/Code/mymaths

Application Trace | Framework Trace | Full Trace

app/controllers/assessment/admin_controller.rb:41:in `class_admin_index'

config/application.rb:116:in `call'

Request

Show env dump

Response

Headers:

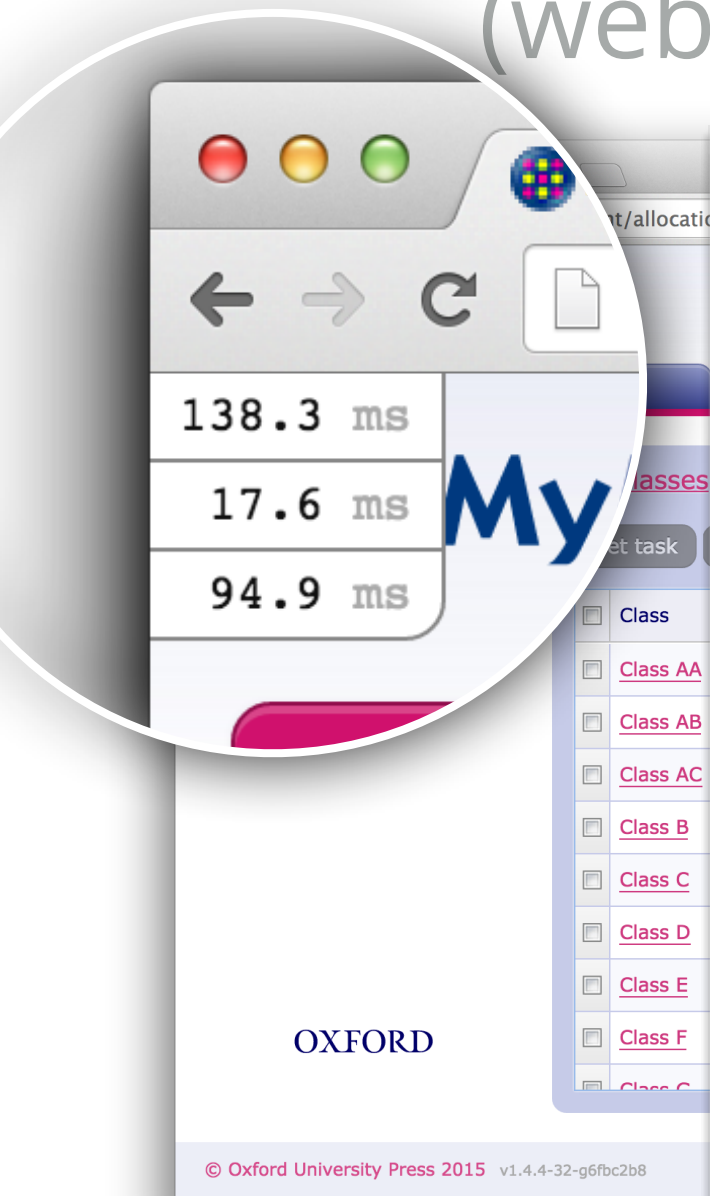
None

Timeout::Error: execution expired:

© Oxford University Press 2015 v1.4.4-32-g6fbc2b8 Help | Contact | News | Privacy | Legal | Terms & Conditions | Cookie Policy

Profiling

(web middleware can help big time!)



	duration (ms)	from start (ms)	query time (ms)	
GET http://localhost:3000/assessment/allo...	18.2	+0.0		
Executing action: allocation_index	63.2	+17.0	2 sql	0.5
Rendering: /assessment/manager/man...	9.3	+80.0		
Rendering: /assessment/manager/ma...	1.6	+85.0		
Rendering: /assessment/manager/m...	10.3	+86.0	3 sql	0.8
Rendering: /assessment/manager/ma...	7.5	+100.0		
Rendering: /assessment/manager/ma...	5.5	+109.0		
Rendering: layouts/application	12.4	+115.0		
Rendering: /shared/secondary/_header	1.6	+118.0		
Rendering: shared/_footer	3.1	+127.0		

[show time with children](#)

0.9 % in sql

	client event	duration (ms)	from start (ms)
Redirect		22.0	+0.0
Response		1.0	+165.0
Dom Content Loaded Event		168.0	+571.0
First Paint Time			+836.0

[share](#)

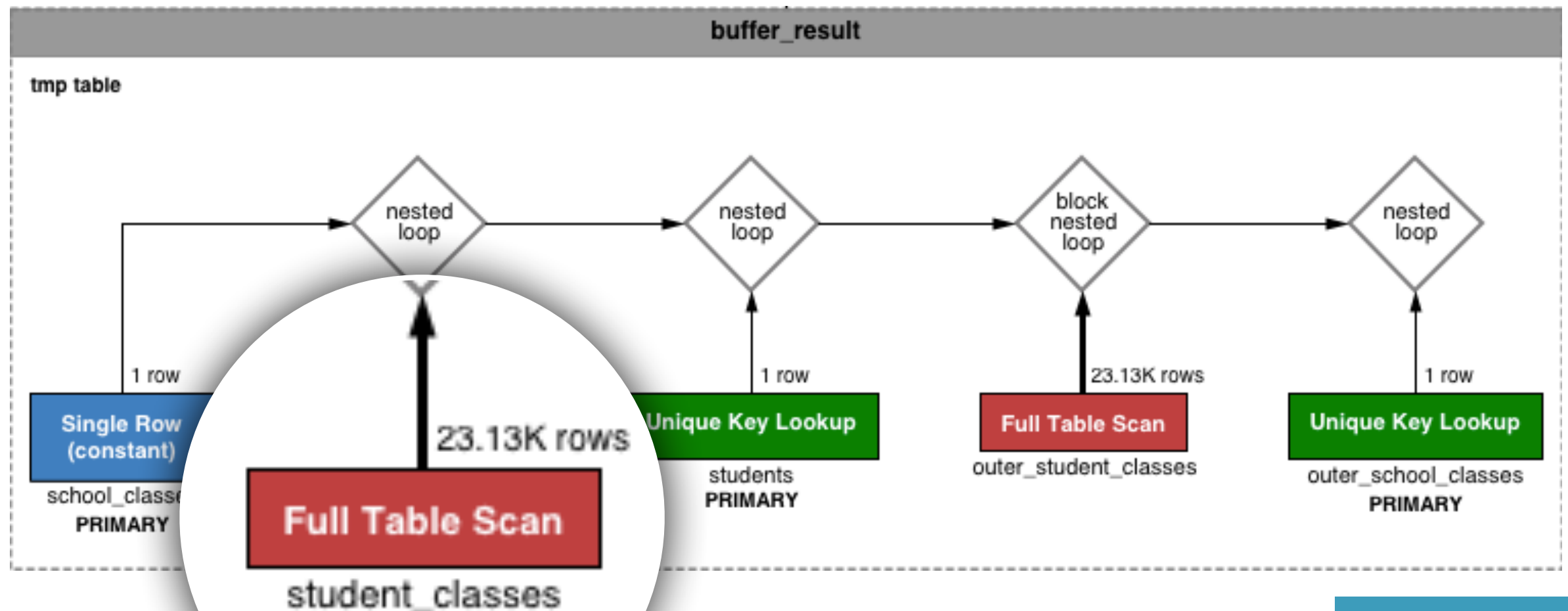
[show trivial](#)



rack-mini-profiler

Pick the low hanging fruit first

(check the database)



Add Database Indexes

(best will in the world, you will have missed some)

```
mysql> EXPLAIN SELECT students.id AS student_id, students.first_name AS forname, students.last_name AS lastName, students.user_name AS loginNumber, GROUP_CONCAT(outer_school_classes.name SEPARATOR ', ') AS class, school_classes.id AS school_class_id, students.password, students.id AS edit FROM 'students' INNER JOIN 'student_classes' ON 'student_classes'. 'student_id' = 'students'. 'id' INNER JOIN 'school_classes' ON 'school_classes'. 'id' = 'student_classes'. 'school_class_id' INNER JOIN 'student_classes' ON 'student_classes'. 'school_class_id' = 'school_classes'. 'id' INNER JOIN 'outer_student_classes' ON 'outer_student_classes'. 'student_id' = 'students'. 'id' INNER JOIN 'school_classes' AS outer_school_classes ON 'outer_school_classes'. 'id' = 'outer_student_classes'. 'school_class_id' WHERE 'students'. 'account_id' = 1028758195 GROUP BY students.id ORDER BY last_name, first_name;
```

id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	school_classes	const	PRIMARY	PRIMARY	4	const	1	Using index
1	SIMPLE	student_classes	ALL	NUL	NUL	NUL	NUL	23130	Using where
1	SIMPLE	students	eq_ref	PRIMARY	PRIMARY	4	mymaths.student_classes.student_id	1	Using index
1	SIMPLE	outer_student_classes	ALL	NUL	NUL	NUL	NUL	23130	Using where
1	SIMPLE	outer_school_classes	eq_ref	PRIMARY	PRIMARY	4	mymaths.outer_student_classes.school_class_id	1	Using index

5 rows in set (0.00 sec)

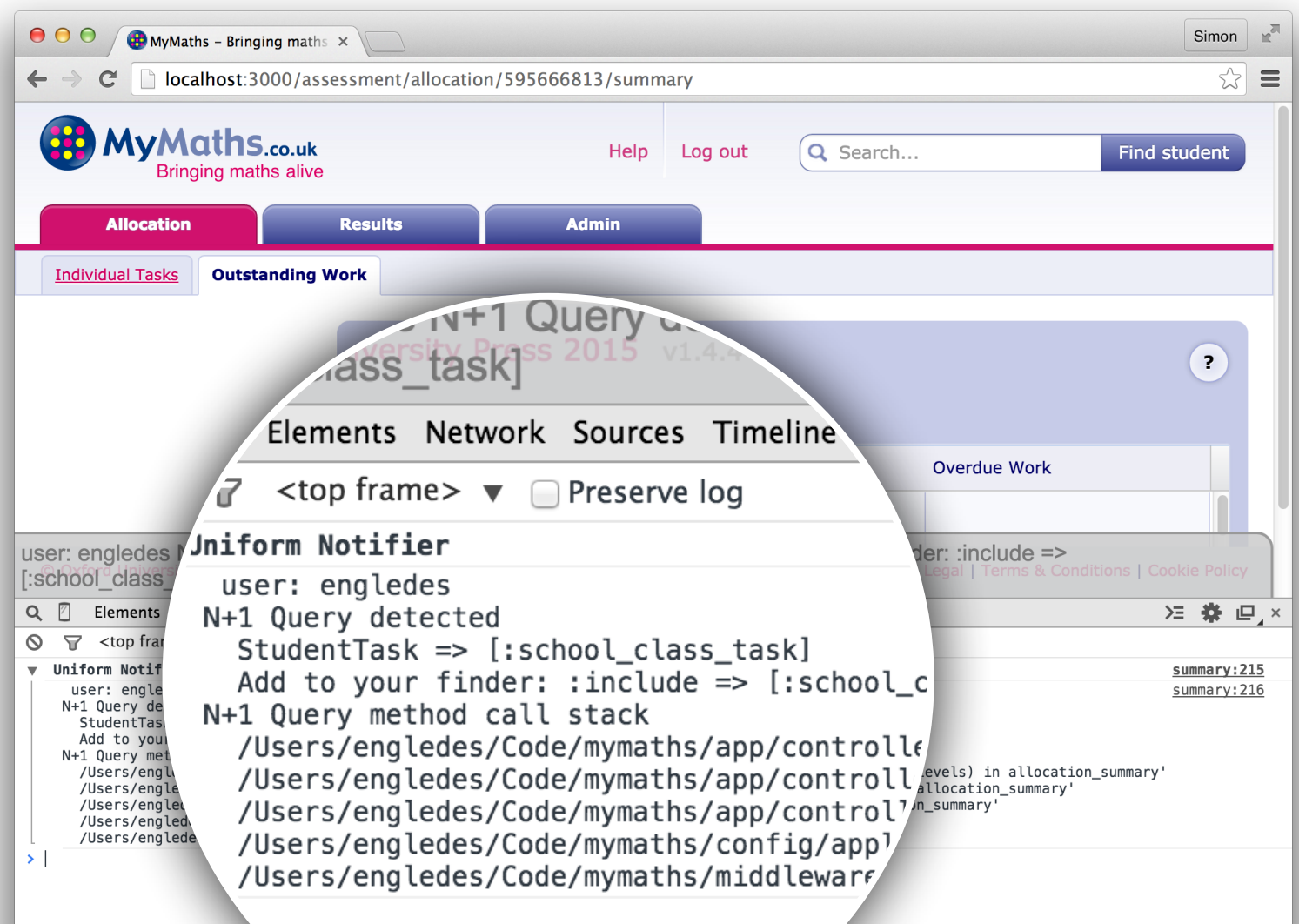
```
mysql> explain SELECT students.id AS student_id, students.first_name AS forname, students.last_name AS lastName, students.user_name AS loginNumber, GROUP_CONCAT(outer_school_classes.name SEPARATOR ', ') AS class, school_classes.id AS school_class_id, students.password, students.id AS edit FROM 'students' INNER JOIN 'student_classes' ON 'student_classes'. 'student_id' = 'students'. 'id' INNER JOIN 'school_classes' ON 'school_classes'. 'id' = 'student_classes'. 'school_class_id' INNER JOIN 'student_classes' AS outer_student_classes ON 'outer_student_classes'. 'student_id' = 'students'. 'id' INNER JOIN 'school_classes' AS outer_school_classes ON 'outer_school_classes'. 'id' = 'outer_student_classes'. 'school_class_id' WHERE 'students'. 'account_id' = 1030095999 AND 'school_classes'. 'id' = 1028758195 GROUP BY students.id ORDER BY last_name, first_name;
```

id	select_type	table	type	possible_keys	key	key_len	ref	rows	Extra
1	SIMPLE	school_classes	const	PRIMARY	PRIMARY	4	const	1	Using index
1	SIMPLE	student_classes	ref	index_student_cl...	index_student_classes_on_school_class_id	5	const	10	Using index
1	SIMPLE	students	eq_ref	PRIMARY, index_st...	PRIMARY	4	mymaths.student_classes.student_id	1	Using index
1	SIMPLE	outer_student_classes	ref	index_student_cl...	index_student_classes_on_student_id_and_school_class_id	5	mymaths.student_classes.student_id	1	Using index
1	SIMPLE	outer_school_classes	eq_ref	PRIMARY	PRIMARY	4	mymaths.outer_student_classes.school_class_id	1	Using index

5 rows in set (0.00 sec)

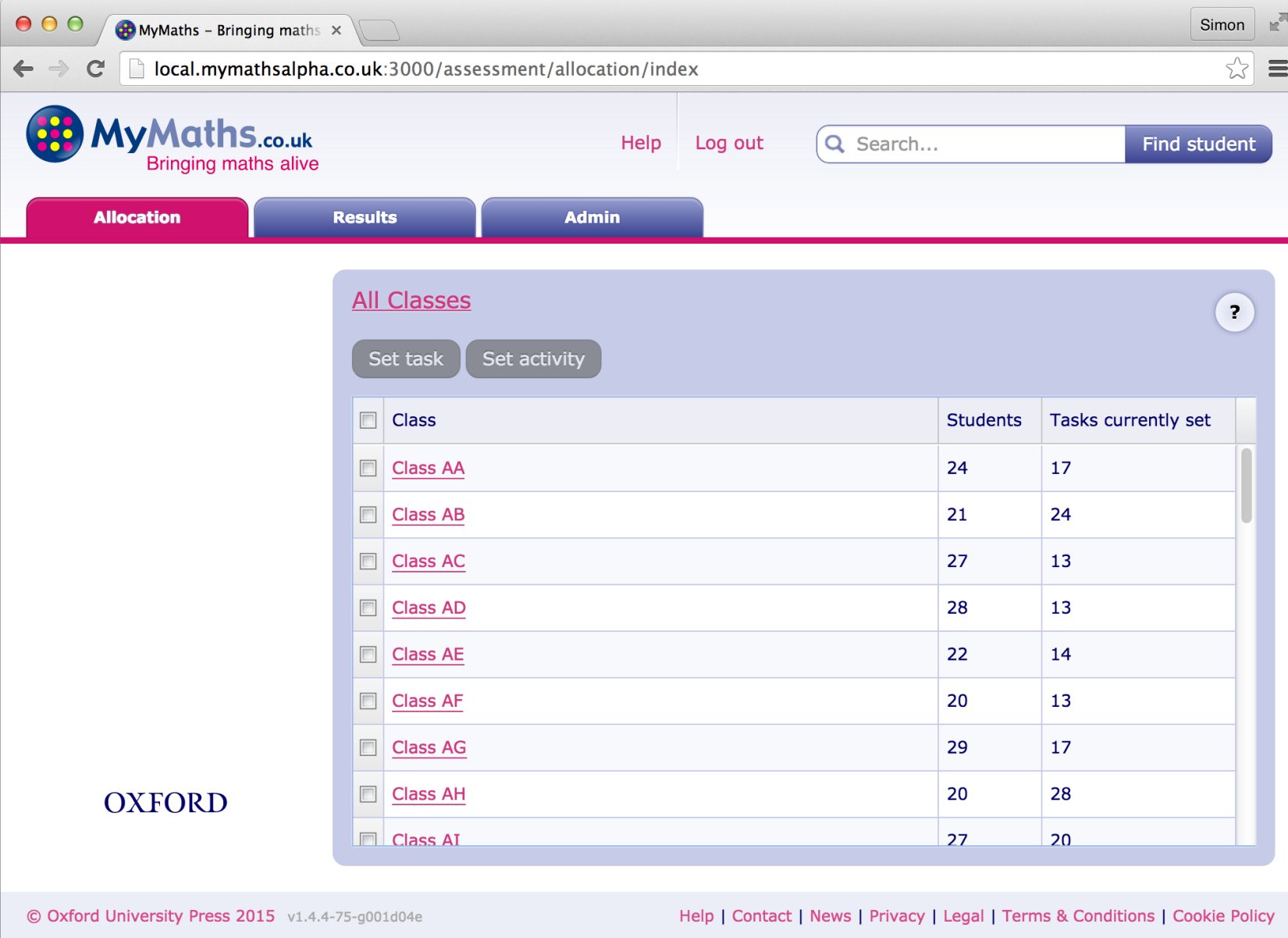
Fix N+1 Queries

(look for loops around your application)



Now the site is loading again

(with some meaty data)



The screenshot shows the MyMaths.co.uk website interface. The browser address bar displays the URL: `local.mymathsalpha.co.uk:3000/assessment/allocation/index`. The page header includes the MyMaths.co.uk logo, a search bar, and navigation links for Help, Log out, and Find student. Below the header, there are three tabs: Allocation (active), Results, and Admin. The main content area is titled 'All Classes' and contains a table with columns for Class, Students, and Tasks currently set. The table lists 10 classes, each with a checkbox, a class name (Class AA through Class AI), a student count, and a task count. The footer includes the Oxford University Press logo and copyright information.

MyMaths.co.uk
Bringing maths alive

Help Log out Search... Find student

Allocation Results Admin

All Classes

Set task Set activity

<input type="checkbox"/>	Class	Students	Tasks currently set
<input type="checkbox"/>	Class AA	24	17
<input type="checkbox"/>	Class AB	21	24
<input type="checkbox"/>	Class AC	27	13
<input type="checkbox"/>	Class AD	28	13
<input type="checkbox"/>	Class AE	22	14
<input type="checkbox"/>	Class AF	20	13
<input type="checkbox"/>	Class AG	29	17
<input type="checkbox"/>	Class AH	20	28
<input type="checkbox"/>	Class AI	27	20

OXFORD

© Oxford University Press 2015 v1.4.4-75-g001d04e

Help | Contact | News | Privacy | Legal | Terms & Conditions | Cookie Policy

Set up and automate your acceptance environment

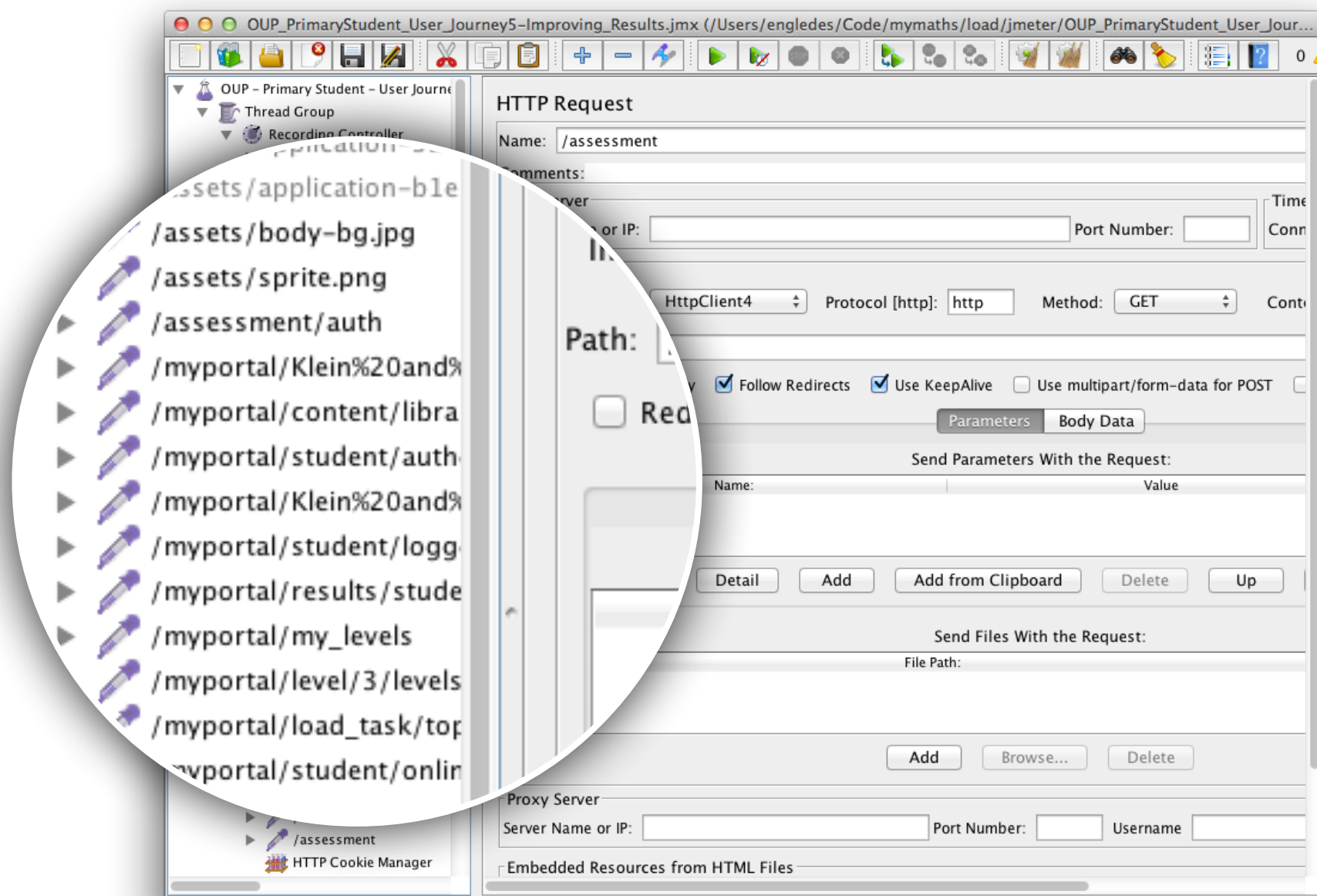
The image displays a screenshot of the AWS Management Console and a CloudFormation template editor. The console window shows the 'Amazon Web Services' dashboard with various service categories like Administration & Security, Application Services, Mobile Services, and Enterprise Applications. A circular callout highlights a section of the CloudFormation template (mymaths-app-elasticache.template) defining an 'mymathsCPUAlarmLow' alarm. The alarm is configured with a 'LessThanThreshold' comparison operator, a 'CPUUtilization' metric, and a threshold of 25. It also references a 'mymathsScaleInPolicy' for alarm actions.

```
402     "Period": "300",
403     "AlarmActions": [ { "Ref": "mymathsScaleInPolicy" } ],
404     "Namespace": "AWS/EC2",
405     "Dimensions": [ {
406       "Name": "AutoScalingGroupName",
407       "Value": { "Ref": "mymathsHostAsg" }
408     } ],
409     "mymathsCPUAlarmLow": {
410       "Type": "AWS::CloudWatch::Alarm",
411       "Properties": {
412         "EvaluationPeriods": "1",
413         "Statistic": "Average",
414         "Threshold": "25",
415         "AlarmDescription": "Alarm when CPU too low",
416         "Period": "300",
417         "AlarmActions": [ { "Ref": "mymathsScaleInPolicy" } ],
418         "Namespace": "AWS/EC2",
419         "Dimensions": [ {
420           "Name": "AutoScalingGroupName",
421           "Value": { "Ref": "mymathsHostAsg" }
422         } ],
423         "ComparisonOperator": "LessThanThreshold",
424         "MetricName": "CPUUtilization"
425       }
426     },
427     "mymathsCPUAlarmLow": {
428       "Type": "AWS::CloudWatch::Alarm",
429       "Properties": {
430         "EvaluationPeriods": "1",
431         "Statistic": "Average",
432         "Threshold": "25",
433         "AlarmDescription": "Alarm when CPU too low",
434         "Period": "300",
435         "AlarmActions": [ { "Ref": "mymathsScaleInPolicy" } ],
436         "Namespace": "AWS/EC2",
437         "Dimensions": [ {
438           "Name": "AutoScalingGroupName",
439           "Value": { "Ref": "mymathsHostAsg" }
440         } ],
441         "ComparisonOperator": "LessThanThreshold",
442         "MetricName": "CPUUtilization"
443       }
444     }
```

amazon
web services™

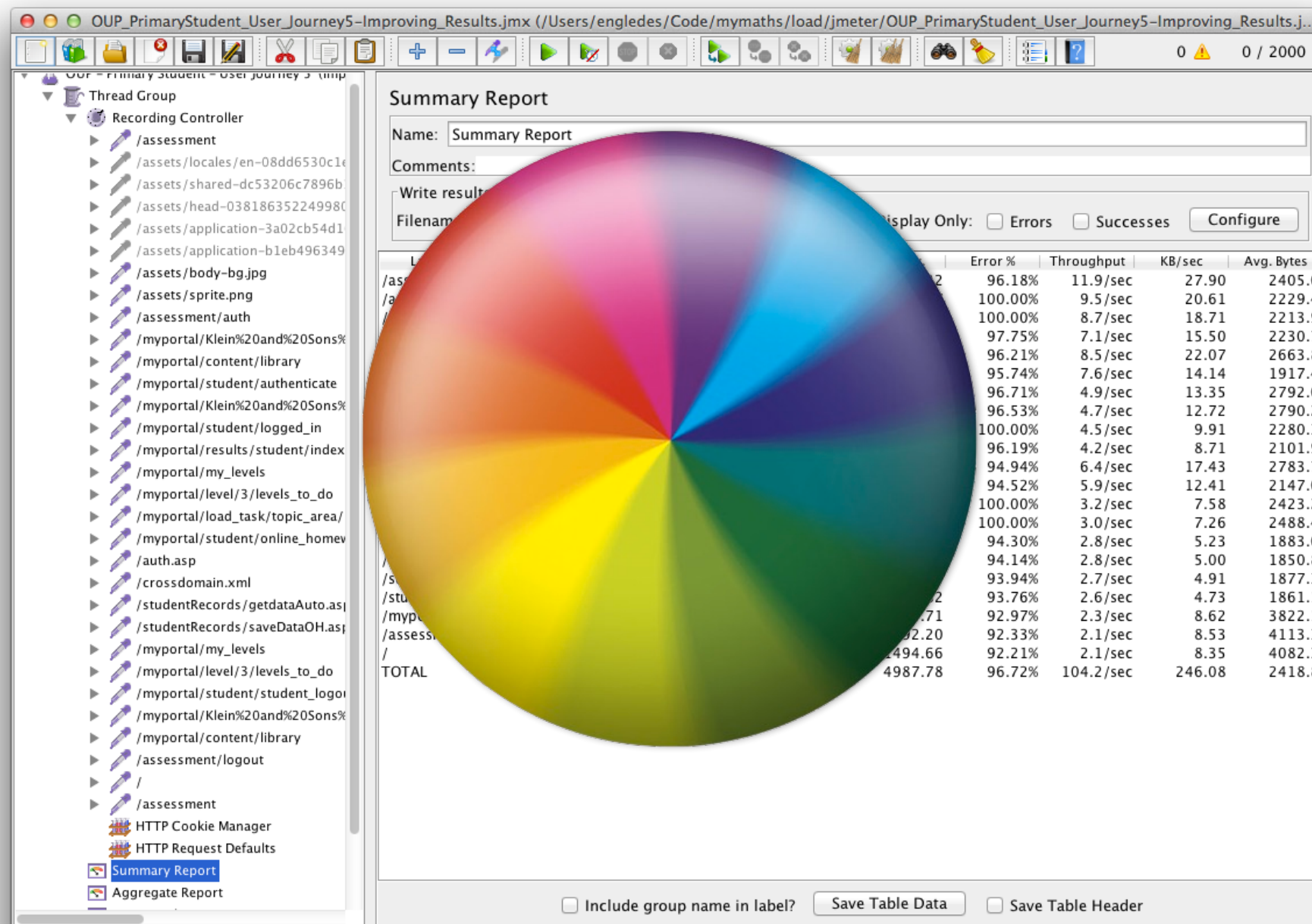
Fire up JMeter

(simulate some user load)



Boom!

(ok, panic a little bit, that should have worked)



Research good production configuration

(turning on threading might help)

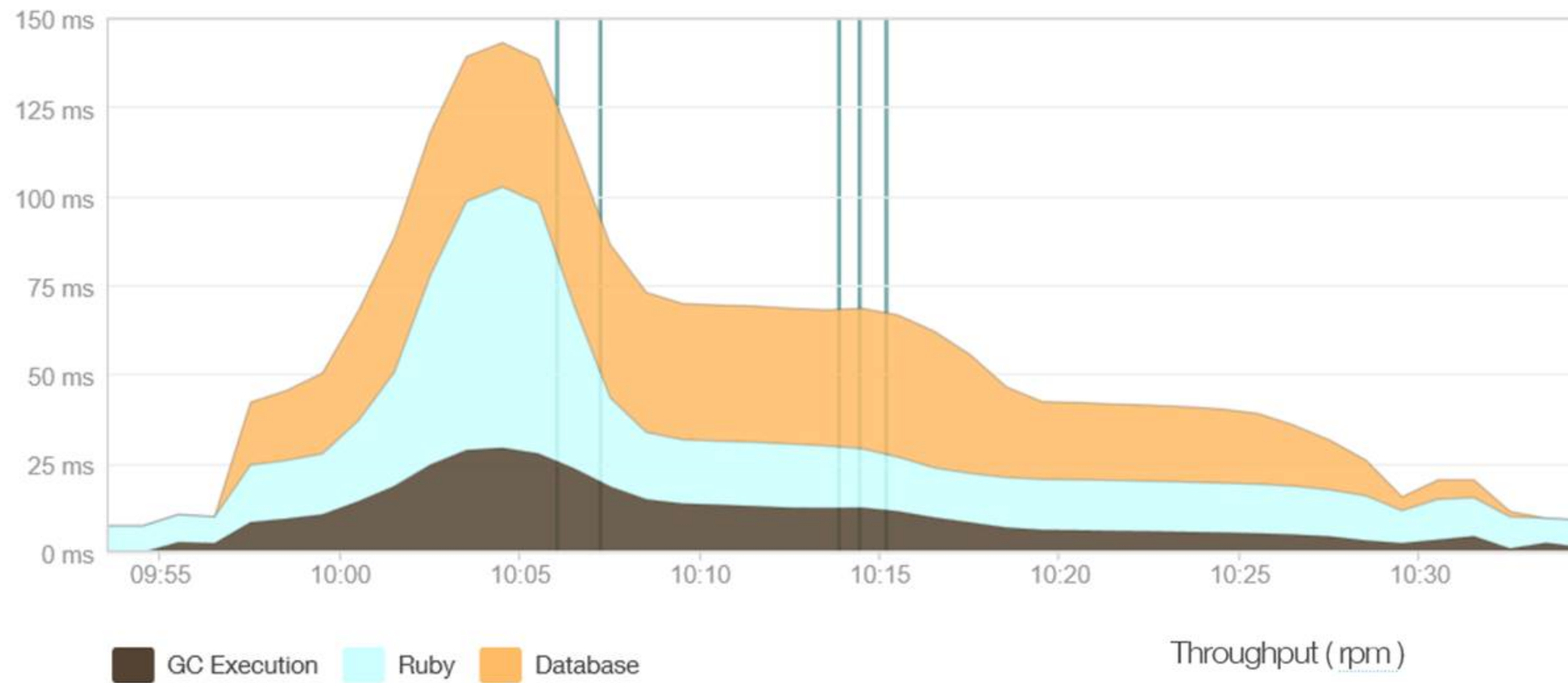
```
20
21 # Defaults to Rails.root.join("public/assets")
22 # config.assets.manifest = YOUR_PATH
23
24 # Specifies the header that your server uses for sending files
25 # config.action_dispatch.x_sendfile_header = "X-Sendfile" # for apache
26 # config.action_dispatch.x_sendfile_header = "X-Accel-Redirect" # for nginx
27
28 # Force all access to the app over SSL, use Strict-Transport-Security, and use secure cookies
29 # config.force_ssl = true
30
31 # See everything in the log (default: :info)
32 # config.log_level = :warn
33
34 # Prepend all log lines with the following tags
35 # config.log_tags = [ :subdomain, :url ]
36
37 # Use a different logger for distributed setups
38 # config.logger = ActiveSupport::TaggedLogging.new(SyslogLogger.new)
39
40 # Use a different cache store in production
41 # config.cache_store = :mem_cache_store
42
43 # Enable serving of images, stylesheets, and JavaScripts from an asset server
44 # config.action_controller.asset_host = "http://assets.example.com"
45
46 # Precompile additional assets (application.js, application.css, and all non-JS/CSS)
47 # config.assets.precompile += %w( search.js )
48
49 # Disable delivery errors, bad email addresses will be ignored
50 # config.action_mailer.raise_delivery_errors = false
51
52 # Enable threaded mode
53 # config.threadsafe! unless $rails_rake_task
54
55 # Enable locale fallbacks for I18n (makes lookups for any locale fall back to
56 # the I18n.default_locale when a translation can not be found)
57 # config.i18n.fallbacks = true
58
```


JMeter Take 2

(better!)

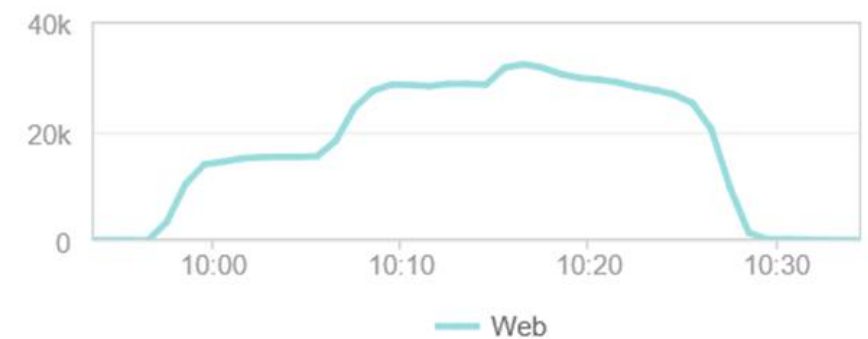
Web transactions response time ▾

64 ms
App server
0 s
Browser



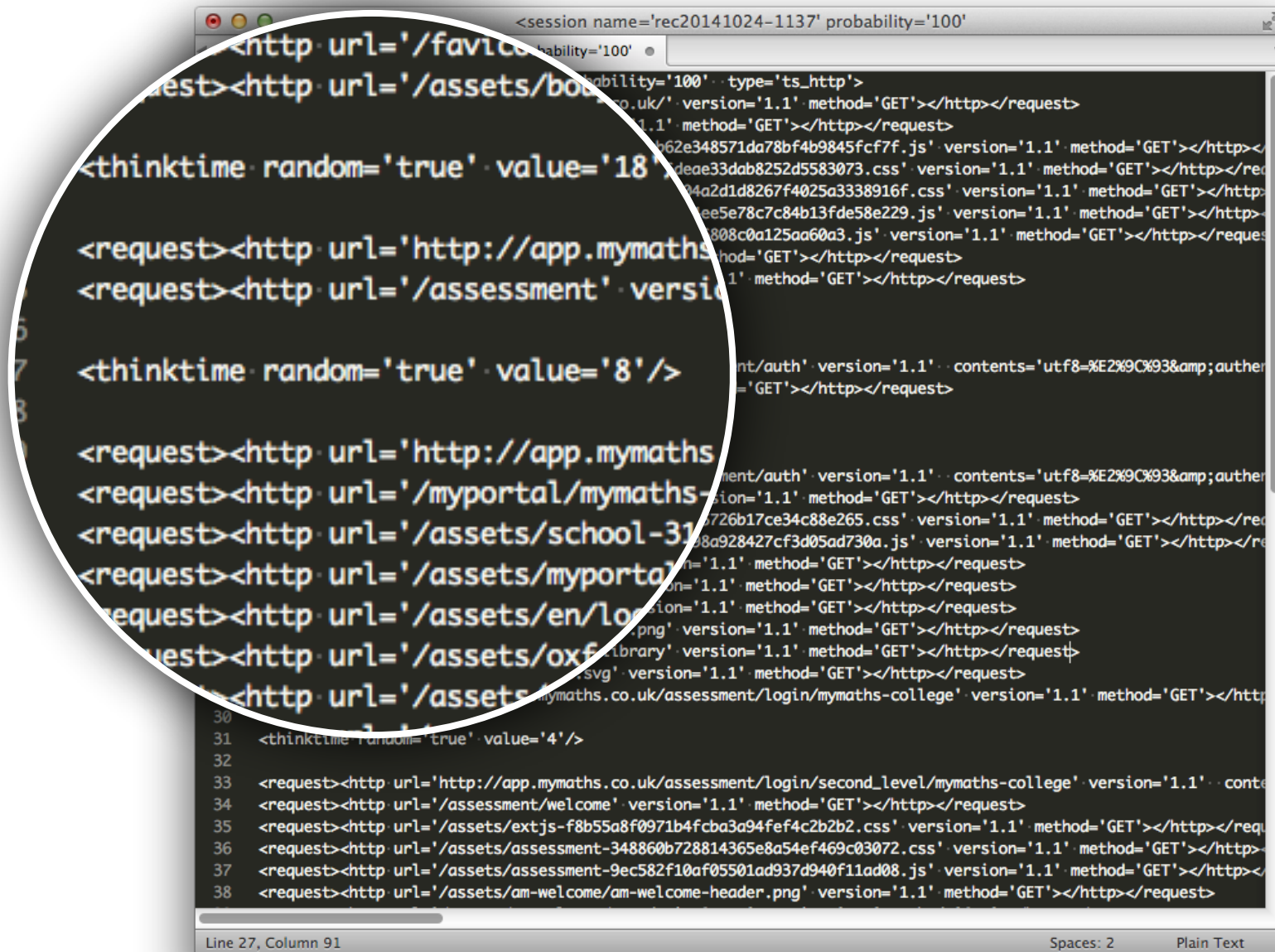
Throughput (rpm)

17,400
Average



If you need to be sure... Tsung

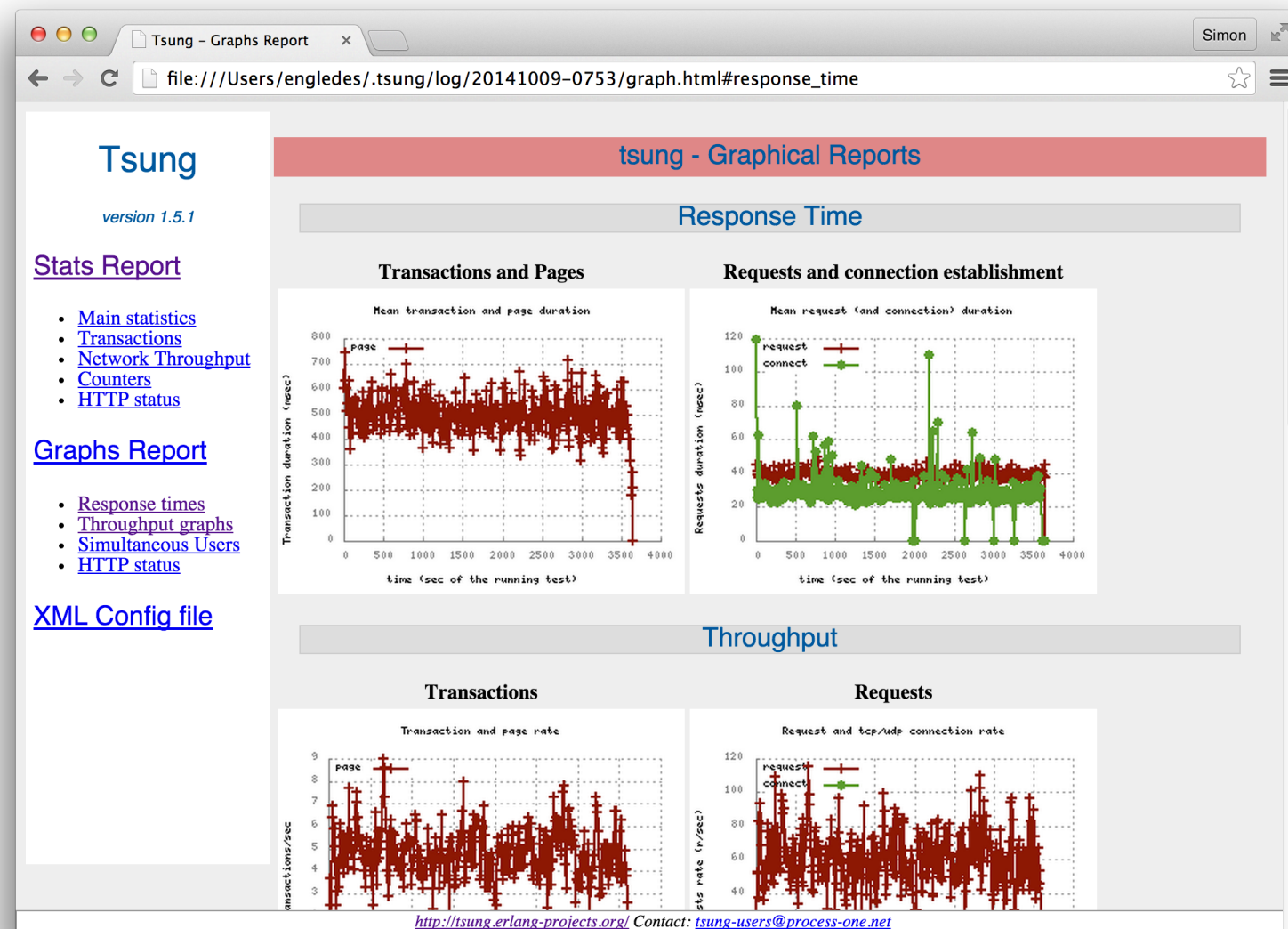
(The big guns)



See also: [siege](#)

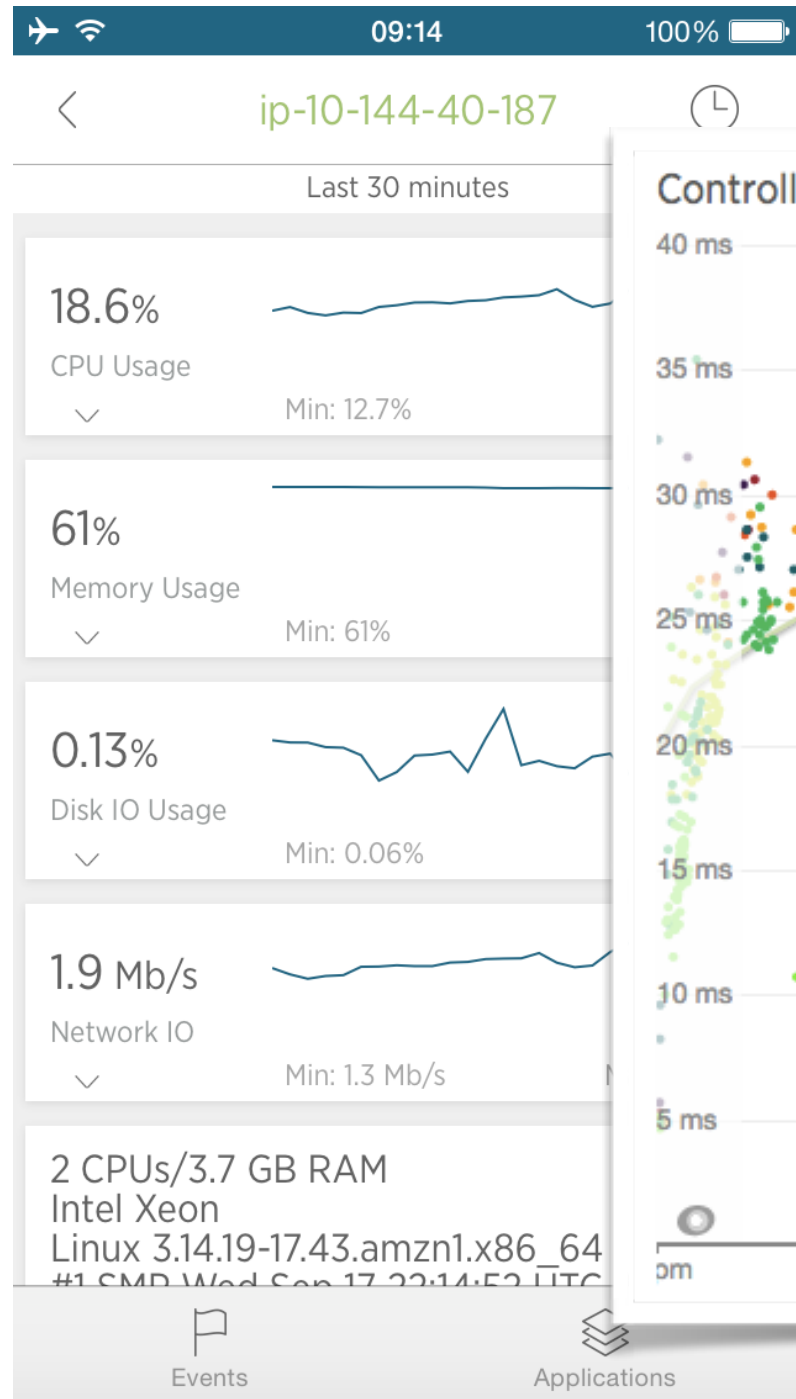
Pour over arcane graphs

(can anyone actually read these?)

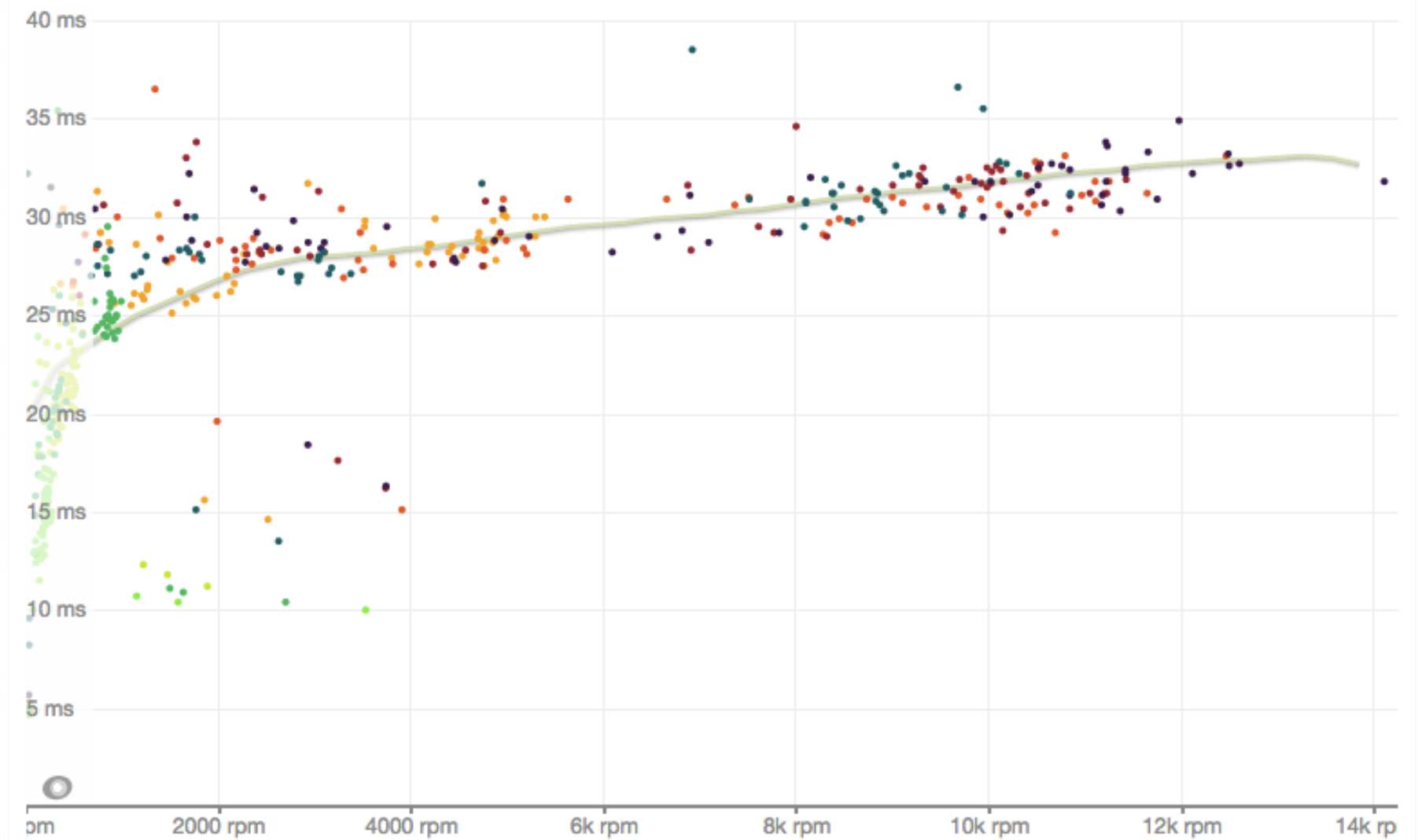


Some Hogwarts grade shit here...

Remote Monitoring



Controller Response Time vs. Throughput



Summary

- Even if you have a small site, generate a lot of fake data to magnify the problems
- Understand and automate your environment
- Every disaster you artificially create is a problem your users never saw - be positive and thorough
- Good monitoring will help you to relax afterwards