



Load Testing of TestWise



sqc

www.sqc.co.uk

SQC Technology Ltd.
Coton Park House, Linton, Swadlincote,
DE12 6RA, UK.

SQC Technology provides software testing, software test management and software test automation services and consultancy. This includes formulating software test strategies, test analysis and development of test suites, test automation and interim management. More information on our advanced capabilities and innovative approaches can be found on the company web-site. www.sqc.co.uk

A first class user experience guarantee; the load testing of TestWise

TestWise a flagship product

Granada Learning supply vital pupil assessment services to schools. Data from assessments helps schools to track a child's progress and shape their education. In 2009 over 600,000 school children took a Cognitive Abilities Test.

TestWise is Granada's flagship digital testing system. Pupils download interactive tests, complete them online and upload their answers for marking. Schools benefit from rapid marking and a rich portfolio of reports turning results into useful information. Usage has grown rapidly as schools move to digital tests. In 2009 schools took 340,000 digital tests during the September to October period.

The availability and error free operation of the TestWise service, especially during the main assessment periods, is critical. Downtime or failures in these periods would have an extremely disruptive effect on the schools that TestWise serves. The availability of TestWise is key to Granada's success with their customers.

Performance challenges

Meeting growing demand September 2009 was a bumper year for TestWise with digital testing reaching record levels. Growth forecasts were between 50% and 100% for 2010. This increasing demand combined with a growing importance as more schools switch from paper to digital to trigger a technology refresh. The aim of the refresh went beyond accommodating one year's growth; it was to establish a sound basis for scaling in subsequent years.

A new platform Granada recognised that to meet these objectives TestWise had to move to a new more sophisticated and scalable platform. Whilst the core technologies remained the same there was a plethora of infrastructure and host changes, a major Oracle version uplift and extensive reconfiguration of the system deployment. The team realised that this was not a risk free exercise and was one that would require rigorous testing; a judgement that proved to be well founded.

Load testing for TestWise

An essential activity The ultimate measure of success is the operation of TestWise under load. The inherent risk involved in the move caused Granada to focus on the need for effective load testing and in December 2009 they began the process of selecting a suitable partner. Granada contacted a variety of potential partners including mainstream tool suppliers and specialist testing consultancies to solicit proposals. Firm proposals from three organisations were evaluated and at the end of January Granada selected SQC to undertake the work.

A comprehensive programme of collaborative testing The project required close collaborative working. Not just for test delivery but also at a higher level to assess the risks, quantify the workload, design the test suite, orchestrate delivery, interpret the results and assist in the resolution of issues. Granada were looking for expertise to help them with all of these aspects. SQC were able to lead the exercise in a way that ensured the work the teams jointly delivered established a high level of confidence in the system. This confidence enabled a decision to be made to migrate the production system onto the new platform.

More than just running tests The successful outcome of the testing was built on a wide set of activities that:

- Analysed historic data to understand the structure and magnitude of the workload
- Identified load related risks
- Catalogued key performance measurements and the associated SLAs.
- Designed a test suite that met the objectives and generated test data to support it.
- Analysed raw test results to provide meaningful information.
- Investigated and diagnosed issues.

An integrated practice Granada's approach was an effective embodiment of good practice. Rather than waiting to the end, until just before go-live, to load test, they recognised the benefits of starting early and of doing continuous testing up until the start of live service. Early in this case meant not only having plenty of

continued over

The right partner

Careful selection For such an important exercise Granada went to great lengths to select a supplier they had confidence in. They were looking for a team that understood the nature of the challenge they would face. Sean Dotchin, the Service Delivery Manager, explains:

"I was set the challenge of subjecting our on-line testing software to Load Testing, to ensure it could withstand forecasted customer usage with additional headroom for the foreseeable future. Given the already aggressive workload on our in-house development and technical teams, I needed to outsource this challenge to a third party.

A rough tender was put together and was issued to three different potential suppliers. All three of which responded with their proposals and associated commercials. Out of the three it was immediately obvious SQC were the preferred third party to do the work."

The developer's choice The development team had a firm opinion on the best way to go. With previous experience of working with the SQC team they were clear who they wanted to work with. Stuart Cairns, the lead developer for the exercise, recalls:

"We required more than just a company that could provide a load testing application, one that was able to go that extra mile by providing advice and expertise on how best to proceed through the exercise. SQC were able to work closely with the in-house staff, effectively becoming a highly efficient virtual team, allowing for diagnosis of problems encountered to be solved by drawing on the diverse knowledge pools both teams provided.

The other aspect which made them stand out from the other companies that were tendering for the work was the attention to detail SQC gave to the initial quotation. They spent a lot of time and effort going through the details of the task they would have to undertake, and produced a thorough outline of their proposal. This allowed us to have confidence that they would be able to provide everything we required to make the project a success."

Deciding factors Whilst technical ability was a key factor in the choice of SQC other important elements included evidence of a sound well thought out approach, confidence in an ability to deliver under pressure, commitment and price. In Granada's eyes SQC won on all of these. Paul Webster, the Operations Director, went so far as to say:

"Out of all the potential partners they were quite simply the most focused. They understood our needs far beyond any of the other partners and went into great detail of how they would approach the load testing. It was clear from the outset that they would bring a lot of much needed experience to the process".

A first class user experience guarantee; the load testing of TestWise

continued...

time; it meant load testing before making key decisions on the architecture and configuration of the platform. Testing supported the project throughout its journey through stages focussing on:

- Major architectural and configuration choices.
- Performance engineering necessary to make the new configuration a viable candidate for further testing.
- Iterative cycles of load testing backed by architectural and performance engineering to address issues and meet the SLAs for performance and reliability.
- Testing how the systems responds to internal failures when under load.
- Ensuring final changes made to introduce September 2010 products did not compromise the systems
- Testing the new platform in its production installation.

Tests that targeted different risks The test suite used different patterns of testing to target different risks; examples include:

- Busy period simulation to provide meaningful response time forecasts.
- Endurance tests to check for creeping or catastrophic failure over time.
- Stress testing to find load related break points.
- Tests to see what happens if internal failures occur under heavy load.

The test campaign

Thorough lab testing Load testing started in earnest in April and ran through to early September. Around two hundred load tests were successfully executed with durations ranging from two to six hours. It is estimated that the system was exposed to around six hundred hours of operation under load. This involved two and a half million simulations of a pupil downloading and completing a test.

Transformation The system went through a transformation, becoming more efficient and more scalable with each step. Early gains saw the demise of the initial choice of load balancer; it was found to be wanting with unpredictable behaviour under heavy load. A mix of tuning changes, a code change and the introduction of additional infrastructure saw the system easily meeting the specified SLAs under the target load.

Resilience Having met performance targets attention then turned to the resilience of the system. Using information gained from triggering failures whilst the system was running under heavy load the Granada team were able to refine their fault tolerance and so reduce the risk that an internal component failure will bring down the service.

The final step The final step on this journey was the installation of the new platform into the production data centre and the migration of the live service onto this platform. Having invested so much effort developing a high performance solution Granada recognised that all of the

investment could be undermined by simple mistakes made when moving to live or by unforeseen influences of the production environment. Unwilling to risk falling at the final hurdle Granada asked SQC to perform further load testing on the production platform. This final safety net caught a number of items that would have caused issues had they lain undetected.

The end result

Success in 2010 In 2010 TestWise experienced its busiest ever period. The fourth week in September saw 80,000 tests taken with the busiest day seeing 18,500 tests. This compared with the busiest week in September 2009 which saw 62,000 tests taken. Throughout this 2010 busy period the quality of service delivered to customers was rock solid.

More relaxed When compared with the same period from 2009 perhaps the most notable thing was the lack of stress experienced by the operations and support people. It was noted by Paul Webster, the Operations Director, that in the 2009 busy period the technical operations team had little time to spare as they tended the system whereas in 2010 things were relaxed enough for time to be spent on other matters not related to keeping the system running.

Room for growth The work done on TestWise has given Granada a system with room for growth; the existing configuration was tested to well beyond the 2010 levels and the scalable architecture should ensure headroom can be maintained by scaling the system as demand grows.

The danger of over simplistic load testing

Granada's needs were well beyond those that could be met by a simplistic remote scripting and shoot load test delivered with little interaction. A number of proposals they received offered this sort of service; minimal client engagement with delivery by a distant team focussing on what they could see on the interface. One offer even limited them to six test runs (more than two hundred were actually executed). Granada recognised the limitations of these offerings and opted for a more comprehensive service that added a lot more value. They recognised the danger of getting superficial load testing where the test team did not really understand how to challenge the system, sacrificed fidelity for simplicity and depended too much on the customer's team for technical leadership.

Scripting and execution with little analysis and low sophistication is now a common offering from some load testing suppliers. Whilst fit for testing simple systems with obvious load characteristics there is a big danger that in many cases, particularly for complex systems and complex workloads, the testing is superficial. The exercises gives the impression that load related risks have been addressed when real-world use soon proves they have not.

Did load testing make a difference for 2010?

How much did load testing contribute to the success of the TestWise technology uplift? Sean Dotchin who holds day to day responsibility for the service, recalls that:

"...during the official load testing, our software and infrastructure needed a reasonable amount of configuration in order to meet the demands of the load testing exercise."

His assessment: of the impact of this was:

"Without doing these configuration changes, and without SQC helping us identify them, our software would not have been able to meet the demands of forecasted customer usage we originally set out to achieve."

In summary the new platform would have failed under load in the busiest and most important period of the year. Paul Webster, the company Operations Director considered what would probably have happened if the load testing had not been done and concluded:

"I don't even want to think about what would have happened if we hadn't carried out such thorough testing."

A first class user experience guarantee; the load testing of TestWise

An ongoing relationship

The future of load testing Granada view load testing as a fundamental part of their development and service management practices. A practice that can both check that planned changes will not have a detrimental effect on the current service and that can forecast how the system will cope with foreseeable growth in usage. Granada need to be able to perform load testing quickly and effectively as and when the need arises.

An ongoing load testing partnership Continuity provides benefits; it ensures availability, eliminates mobilisation costs, reduced lead times and permits continuous improvement. At a basic level it provides better value for money. Recognising this Granada followed the success of the initial TestWise load testing project by establishing a two year partnership arrangement under which SQC maintain and operate a load testing capability for Granada.

This arrangement provides Granada with ready cost effective access to a professional load testing capability, one that will help to guarantee the integrity of Granada's online services in the years to come. No longer is load testing viewed as a bolt on item to be done under exceptional circumstances instead it is part of the normal delivery practices.

Smarter load testing practices

Prescheduled workloads Many work loads, particularly web-loads, are bursty. They vary in magnitude over multiple time scales; wave superimposed on wave. Peak loads can be ten times the mean at any level of granularity. To obtain representative results load testing should, mimic these arrival patterns.

Most load testing tools operate on a cyclic basis. Each virtual user resumes activity after a regular delay. In some cases a randomly generated delay can be used but its distribution is generally uniform. Attempts can be made to use this random element to vary the load but these techniques do not yield good simulations of real loads; they tend to produce loads that vary but that do not vary in the way that real loads do.

There is also a second issue with this approach. The load is randomly generated each time a run occurs. The load cannot be repeated; each run is different. Now suppose a problem occurs during a run, it may be that it cannot be reproduced as any repeat run has a different load pattern. Tests that cannot be repeated are not very effective tests.

The technique used to address these issues in TestWise load testing was to use an offline scheduling program to generate a real world like schedule of events. The tests were implemented within a framework that could process the schedule data and use it to control exactly when visitors interacted with the system. The benefits of this were.

- The generated loads mimic real loads accurately.
- The load structure can be checked before execution.
- Tests became repeatable and so did system failures and performance issues..

Interleaving Typically virtual users spend much of their time doing nothing; sleeping to simulate user delays. Servers do not work like this; instead threads work in pools each serving different connections as work arrives; this avoids threads sitting around waiting. Applying this pattern to load testing for TestWise each virtual user performs operations for many concurrent visitors. We easily simulate seven thousand concurrent visitors with one thousand virtual users. We reduce the number of virtual users required and this in turn reduces the amount of equipment needed to generate the load.

Testing from the Cloud

When challenged to test the production platform SQC used the Cloud. It provided a large scale load generation capability with high bandwidth connectivity into the internet. The Amazon EC2 platform proved an ideal place to host load generators that could be activated and deactivated on demand. Testing from the Cloud proved a highly effective method for testing the production platform.

Hosting load generation in this way is highly scalable and ensures that the available bandwidth grows alongside any growth in load generation capacity. Access to the effectively limitless amount of virtual infrastructure available means that equipment limits are no longer a significant issue when load testing. When testing large scale platforms via their production front ends testing from the Cloud is now a prime candidate.

Working with SQC

Load testing on TestWise was a significant intellectual, technical and delivery challenge. It involved a lot of hard work, commitment and innovation. The quality and level of service provided by SQC and the efforts put into helping Granada achieve its objectives were clearly recognised by Granada.

Paul Webster, the Operations Director, indicated his willingness to recommend SQC to other potential customers as a result of his team's experience of working with SQC on this exercise. ::

"The level of service was second to none. They provided a very professional service throughout, working very closely with the technical team. I would recommend SQC to anyone."

Sean Dotchin, the TestWise Service Delivery Manager, is the man at the sharp end. He is responsible for the day to day delivery of the TestWise service to customers and was tasked with overseeing this testing. His views on the way the exercise went are:

"Throughout the load testing, SQC provided a service which exceeded any expectation we had when we initially started the exercise. Everything from the SQC Team individuals, the documentation, the support, the advice, the approach, far exceeded what any other potential load testing partner could have achieved, and we would ultimately have proceeded to have serious technical issues with our software had we not undertaken the exercise."

Recognition of the success of the engagement was not just at a formal delivery level. An effective, collaborative and friendly working relationship quickly sprang up and persisted throughout the most difficult and stressful stages of the delivery. Sean went on to say:

"As a result of the exceptional experience with SQC, the on-going relationship is becoming more concrete, as we look to partner with SQC for future load testing and wider IT consultancy services."

His concluding comments were::

"Would I recommend SQC to other customers? Without a shadow of a doubt, probably the most recommended and trusted partners I have the pleasure to work with in over twenty years in the industry."

More Information?

To find out more about the things we can do to help you to manage your performance and scalability risk please send an email to at enquiry@sqc.co.uk and we will contact to discuss what we do and what you need.