Blaise 4.8.4 Web Form Load and Performance Testing

Author(s): Oleg Volguine, Presenter(s): Helen Robson, Lane Masterton

Organization: Australian Bureau of Statistics

1. Abstract

This is a technical paper on load and performance testing of Blaise 4.8.4. The Australian Bureau of Statistics (ABS) has recently introduced web surveys for its collections using Blaise Internet. In order to ensure the quality of provider experience the ABS has undertaken both internal and external load and performance testing. This paper details the investigations undertaken, including the methodology, approach, and test strategies utilised, as well as the results and findings that were obtained.

2. Background

ABS has made a strategic decision to use Blaise 4.8.4 (Blaise IS) for all web form development. The first household collection released on Blaise IS in December 2012 was the Monthly Population Survey (MPS). Migration of business surveys to Blaise IS commenced in December 2012 and will be completed in July 2014. Throughout the June and September quarters for 2013, the eCollection platform is expected to support up to 142,000 eForm submissions across many business and household survey collections.

The ABS has undertaken load and performance testing in order to ensure a stable and responsive online survey respondent experience. The testing for Blaise eCollect system has been carried out by ABS staff, and through an engagement with an external load and performance testing partner, to establish an understanding of Blaise IS capabilities, optimal infrastructure configuration, scalability options and how this solution supports the anticipated business outcomes for surveys going live in June 2013 and beyond.

3. Purpose of Testing

3.1. Purpose

The purpose of the test program was to evaluate the capacity of the Blaise 4.8.4 eCollect platform to process the expected production load levels for June 2013 ABS eForms. Specifically the goals of this program were designed to:

- Assess the response times of key transactions under production load, such as login (including both authentication and authorisation), completion of survey data, and survey submission.

- Evaluate the reliability of the Blaise 4.8.4 infrastructure while processing production levels of load over an extended period.

- Assess system performance for different end user network speeds, 56kbps, 64kbps, 512kbps and 2048kbps.

- Identify any bottlenecks impacting system performance and highlight options for resolution.

- Facilitate an initial round of system performance diagnosis and optimisation.

- Provide key learnings for further performance and load testing and to support future system optimisation.

3.2. Load and Performance Targets and Service Level Agreements (SLAs)

In order to ensure a stable, reliable and responsive provider experience, the following SLAs and key performance metrics were targeted:

- 5 seconds for at least 90% of all other transactions (page to page transactions).

- 15 seconds for at least 90% of login (authentication and authorisation) transactions.

- Stable and responsive system behaviour over time:

  That is, no system performance degradation over time such as memory leaks or excessive use of hard disk space or system instability. The business requirement is that the system is operational 24 hours per day and 7 days a week.

4. Test Strategy

The test strategy consisted of a suite of tests designed to assess system performance, reliability and responsiveness. The full suite of tests and their purpose is described in this section.

4.1. Normal (Average) Capacity Test

The purpose of this test was to assess the system's ability to cope with the average production level load expected on a peak day. Average load was defined as the load that the system was expected to cope with on its busiest day. The average value was derived by taking the total number of transactions on the busiest day and averaging over the time that the system was expected to be used. For the Blaise eCollect platform this value was derived by looking at the historical number of paper form submissions for all surveys that were expected to be migrated to eForms by June 2013 and combining the number of submissions on the peak day for all surveys over the full enumeration period. The total was then averaged over the number of hours that the system was expected to be operational on any day which was 11.

4.2. Peak (Absolute) Capacity Test

The purpose of this test was to assess the system's ability to cope with the absolute maximum load expected on a peak day. For example, on a peak day, for a period of time the system may experience a load that is much higher than the average peak day load. For the purpose of testing the Blaise eCollect system this value was based on historical information from previous online surveys and from the 2006 and 2011 Australian Population eCensus experiences. This value was set at 1.7 times the expected average capacity load.

4.3. Endurance Test

The purpose of this test was to assess the system's ability to cope with a sustained load over a prolonged period of time, the test was set for 8 continuous hours. For this test, the system must remain stable and responsive without performance degradation over time such as increased usage of memory (memory leaks), excessive use of hard disk space, system instability or failure.

4.4. Stress Test

The purpose of this test was to assess the limits of the system's performance under a given hardware and system configuration. The test was designed to push the system's limits far beyond the expected production levels of load. For the Blaise eCollect system the levels of stress test load were set based on the results of the peak capacity tests and were double the load set for the peak (absolute) capacity test.

4.5. Data Extraction Test

The purpose of this test was to assess the system's ability to cope with the production level load expected on a peak day while data was being extracted from the live production database and loaded into the back-end ABS systems. The business requirement for this was that the system must allow for regular data extraction and loading of submitted surveys into back-end processing systems (one every hour), while continuing operating normally. The regular loading of data into ABS systems provides up-to-date management information for further follow-up with survey providers.

5. Load Modelling and Test Methodology

5.1. Modelling the Expected System Load

The modelling of the expected system load was based on estimating the maximum number of respondents expected to use the system at any particular time. The aim of this approach was to estimate the number of concurrent users and total survey submissions per hour. The transactions per hour target could then be used to derive more fine-grained transaction rates if needed, such as transactions per minute or transactions per second.

The modelling for the expected system load was based on existing metrics from the ABS paper based business survey returns and interviewer based household forms returns. For the round of load testing undertaken for June 2013, all available information on respondent return rates for all surveys that were going to be offered as eForms in June were analysed and combined to derive the total number of returns per day. The largest number of combined daily survey returns expected on any given day was 3,938. This number was then used to derive the average peak hourly rate of survey submissions, by dividing 3,938 by the number of hours in a day that respondents could use the system. For example, while the eCollect platform operates 24 hours, 7 days a week, the system is only significantly utilised for 11 hours, between the hours of 8am to 7pm. This figure was based on the Blaise eCollect survey metrics from the March 2013 quarter.

In addition to the average peak hourly rate, the absolute peak hourly rate was also used in order to account for a load that was higher than the expected average hourly load. For example, on a peak day, for a period

of time the system could experience a load that was much higher than the average peak day load. For the purpose of testing the Blaise eCollect system this value was based on historical information from previous online surveys and from the 2006 and 2011 Australian Population eCensus experiences. This value was set at 1.7 times the expected average capacity load.

The average peak and absolute peak targets were derived as:

Average Peak hourly rate: 3,938/11 = 358 survey submissions per hour.

Absolute Peak hourly rate: 1.7 x 358 = 607 survey submissions per hour.

The peak hourly rates were then used to derive the expected number of concurrent users required to achieve the target number of survey submissions an hour. This approach was based on overall target submissions an hour and the average time it took to complete the survey. It was also based around the assumption that not all users would start and complete their surveys at the same time and in step with each other. For example, if the target hourly rate was 360 survey submissions then the users would have to be working through the survey at a rate of 6 surveys per minute to reach this target. Also, if the survey took 20 minutes to complete and submit, then in 20 minutes, 20x6 =120 submissions would be expected. Therefore, the required minimum number of independent users or concurrent users required to complete their surveys at a rate of 6 submissions per minute would have to be 120. This was the number of users at any point in time throughout the peak hour who would be concurrently working to complete and submit their survey.

Each of the survey tests was then configured to repeat the necessary number of times for each concurrent user in order to achieve the target survey submissions per hour figure. In the example given above, this was 3 iterations per user in one hour.

5.2. Test Methodology

For each survey selected for testing, a specific scenario was designed in detail and documented. This scenario included the most common sequence of questions answered by most survey respondents for that survey. This was done to ensure that a reasonably realistic load was placed on the system.

The scenario containing the question and answer pattern was then recorded for each survey using HP Performance Centre Load Runner version 11. This tool is an industry standard solution for load and performance testing. This produces an automated script, used to simulate the behaviour of the end user from the point of view of their interactions with the web based eCollect system. For example, a typical scenario that may be expected of users is to login, complete a specific series of questions relevant to the sample group, and then submit the survey.

Out of twelve ABS surveys scheduled for eForm migration in June 2013, a combination of representative surveys from Economic (Business) and Population (Household) programs were selected for their characteristics in terms of providing valuable information about system behaviour under load. These characteristics included the sample population size, survey structure and the number of question fields in the survey. The number of simulated users for selected surveys was scaled upwards to account for those surveys that were not tested. The list of selected surveys for testing is detailed in the Appendix, Table 3.

The recorded scripts were then run using the HP Performance Centre load generation platform, which simulated the targeted number of concurrent users and survey submissions per hour in terms of their interactions over HTTP/HTTPS web protocols with the Blaise eCollect system.

The performance of infrastructure components of the eCollect solution such as memory, CPU, disk space and network bandwidth utilisation were monitored throughout the duration of each of the tests, and the results analysed using the analysis toolset available in HP Performance Centre solution.

6. ABS Blaise 4.8.4 eCollect Solution Architecture

The following diagram, Figure 7, shows the architectural solution overview for the Blaise eCollect platform. The diagram depicts all major solution components as well as major communication flows.

A description of each of the components is presented in the next section.
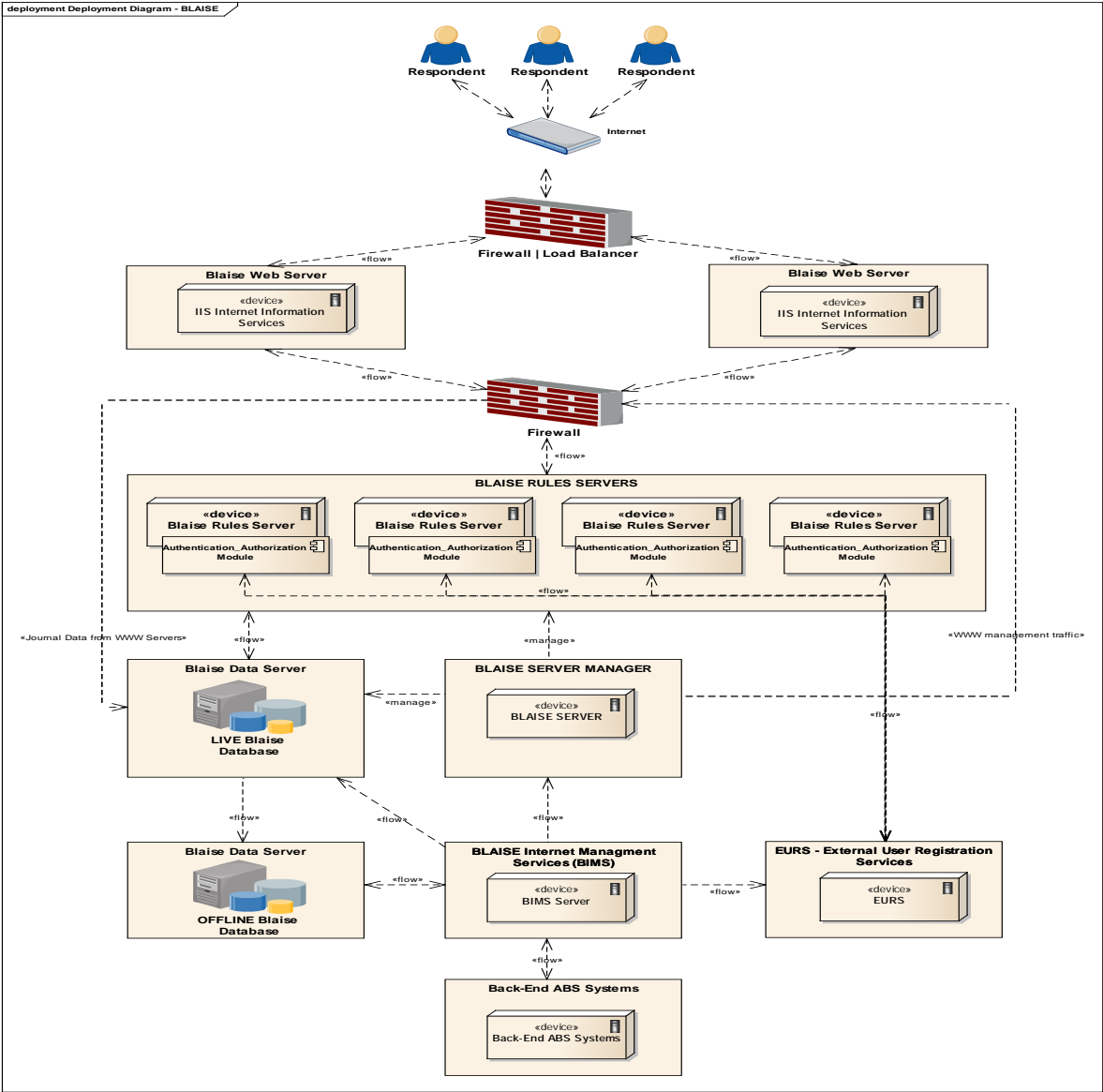

Figure 1 Blaise 4.8.4 eCollect Solution Architecture

6.1. ABS Blaise 4.8.4 eCollect Solution Components

The ABS Blaise 4.8.4 eCollect solution consists of a number of infrastructure components. It is based around the Blaise IS product and it is built using a single Blaise Park. A Blaise Park is a combination of one or more Blaise Web Servers, one or more Blaise Rules Servers, a Blaise Data Server and a Blaise Management Server. In addition the Blaise eCollect platform has a number of ABS specific components that allow for integration with ABS processing systems.

Blaise Web Servers

The Blaise Web Servers provide the presentation layer for Blaise Internet. Web Servers interact with the Blaise Rules Servers when a user navigates to the next survey page or submits a survey.

Blaise Rules Servers

The Blaise Rules Servers are responsible for executing the business logic for web surveys. For example, Rules Servers determine the next sequence of questions to be presented based on the answers received.

Blaise Management Server

The Blaise Management Server coordinates the Blaise Server Park and provides functionality for deployment and management of surveys deployed to Blaise Internet.

Blaise Data Server (Live Database)

The Blaise Data Server (Live Database) stores collected provider data for the duration of a survey.

Blaise Data Server (Offline Database)

The Offline Database stores provider response data once it has been submitted. This component is used to mitigate security risks associated with storing provider data in the Live Database.

Blaise Internet Management Services (BIMS)

The Blaise Internet Management Services (BIMS) component is an ABS built Web Service layer for eCollect that provides interfaces to survey lifecycle management. Typical operations facilitated by BIMS include, initialisation of a survey, retrieval of collection processing status and retrieval of provider response data (data extraction). BIMS integrates Blaise Internet with back-end ABS processing systems through the use of the Blaise API.

External User Registration Services (EURS)

External User Registration Services (EURS) is an ABS developed system that is used to manage provider credentials for authentication purposes.

Authentication and Authorisation Module

The Authentication and Authorisation Module is an ABS developed component. The primary purpose of this module is to provide authentication and authorisation functionality for survey respondents. It is used

in conjunction with the ABS External User Registration Services. The module is implemented as a DLL and is installed on each of the Blaise Rules Servers. This module is referenced and called through the use of Blaise 'ALIEN' procedure calls in Blaise instruments.

7. Test Environment

7.1. Setup

The test environment was setup as shown in the solution architecture diagram Figure 1. The server hardware components in the test setup were configured as per the details in Table 1. All servers in this environment were virtual and not physical machines. Testing included all valid network components, including firewall, routers/switches and load balancer devices.

Table 1 Test Environment Configuration

Blaise Park Component	Operating System	Software	Hardware Specification
Blaise Web Server 2 Servers	Windows Server 2008 R2	Blaise 4.8.4.1767 Microsoft Internet Information Services (IIS 7)	4 x CPUs @ 2.7Ghz Intel Xeon E5-26800 [*] 4GB RAM
Blaise Rules Server 4 Servers	Windows Server 2008 R2	Blaise 4.8.4.1767	2x 4 CPUs @ 2.93Ghz Intel Xeon X5570 2x 4 CPUs @ 2.7Ghz Intel Xeon E5-26800 4GB RAM
Blaise Data Server 1 Live DB Server 1 Offline DB Server	Windows Server 2008 R2	Blaise 4.8.4.1767	4 CPUs @ 2.93Ghz Intel Xeon X5570 4GB RAM 2 CPUs @ 2.93Ghz CPU Intel Xeon X5570 4GB RAM
Blaise Management Server 1 Server	Windows Server 2008 R2	Blaise 4.8.4.1767	2 CPUs @ 2.93Ghz CPU Intel Xeon X5570 2GB RAM
BIMS Server 1 Server	Windows Server 2008 R2	Blaise 4.8.4.1767 Microsoft Internet Information Services (IIS 7)	2 CPUs @ 2.93Ghz CPU Intel Xeon X5570 2GB RAM

*The number of CPUs on the Web Servers was upgraded from 2 to 4 CPUs based on results of Stress Test 1 – Section 8.4.

7.2. Load Generation

Load was applied to the Blaise eCollect system by load generators external to the ABS infrastructure environment. This was done to ensure that all ABS eCollect infrastructure components, including load balancers, firewalls and routers were tested.

7.3. Test Monitoring

All Blaise eCollect components were monitored throughout test runs by using the ABS PG3 Tool, which automatically monitors server resource usage (CPU, memory, disk and network performance). Additionally, Windows Performance Monitor tool and verification of system logs were used to assess system behaviour under load.

8. Test Results

This section describes the tests that were conducted and the results of those tests.

This section does not detail every single test that was undertaken, as multiple iterations of tests were often undertaken, particularly if any issues were encountered during test runs. For clarity, only final and significant test results are included, in order to highlight Blaise eCollect performance characteristics and issues which were encountered. In addition to this, analysis of Blaise scalability and performance as well as issues and challenges are included in Section 10 and Section 11 respectively.

8.1. Normal (Average) Capacity Test Results

Test Parameters: 127 Concurrent Virtual users for 2 hours, target 397 submissions an hour.

Execution: 11/05/2013 between 11:25:03 -13:52:01

Objective: Normal Load was applied to Blaise eCollect system using various network speeds and benchmark end user response times for MPS, QBIS and REACS surveys.

The following network speeds were applied: 56Kbps, 64Kbps, 512Kbps and 2048Kbps.

Key Observations:

- The total number of survey forms submitted was 795 and it was as per transaction rate of 397 submissions an hour targeted for peak load test.

- There were no errors seen throughout the execution of run.

- Response times for 90% of the transactions for submission of survey pages at 512Kbps and 2048Kbps are within acceptable SLA of 5 secs.

- Response time for 90% Login module transactions and click survey transactions at 512Kbps and 2048Kbps is within the SLA of 15 secs.

- Response times for 90% of the transactions for submission of survey pages at 56Kbps and 64Kbps exceed SLAs of 5 seconds, and are in the range of 10-20 seconds, and as high as 40 seconds.

- Response times for 90% of the Login module transactions and click survey transactions at 56Kbps and 64Kbps exceed 15 seconds and are as high as 80 seconds.

- CPU utilization on Web Servers was averaging 35%, 10% on the rule servers, and less than10% on the Database server.

8.2. Endurance Test

Test Parameters: 127 Concurrent Virtual users for 8 hours, target 397 submissions an hour.

Executed: 11/05/2013 18:40:52 - 03:49:02

Objective: The objective of the test was to apply normal load to Blaise eCollect system for 8 hours using various network speeds and benchmark end user response times for MPS, QBIS and REACS surveys to verify if system can handle load for prolonged period.

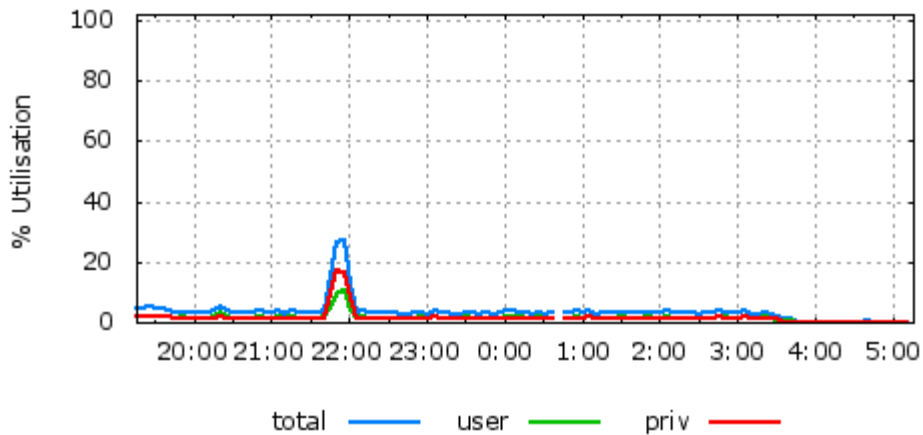The following network speeds were applied: 56Kbps, 64Kbps, 512Kbps and 2048Kbps.

Key Observations

- The total number of survey forms submitted was 3,231 and it was as per transaction rate of 397 submissions an hour targeted for peak load test.

- No errors and transaction failures encountered throughout the duration of the run.

- There was no degradation in response times under load over the 8 hour period.

- No memory leaks were encountered during the total duration of the run.

- Response times for 90% of the transactions for submission of survey pages at 512Kbps and 2048Kbps are within acceptable SLA of 5 secs.

- Response time for 90% Login module transactions and click survey transactions at 512Kbps and 2048Kbps is within the SLA of 15 secs.

- Response times for 90% of the transactions for submission of survey pages at 56Kbps and 64Kbps exceed SLAs of 5 seconds, and are in the range of 10-20 seconds, and as high as 40 seconds.

- Response times for 90% of the Login module transactions and click survey transactions at 56Kbps and 64Kbps exceed 15 seconds and are as high as 80 seconds.

- CPU utilization on Web Servers was averaging 35%, 10% on the rule servers, and less than10% on the Database server.

8.3. Endurance Test Graphs

CPU Average Utilisation – total   11May13 – 12May13

Web Server 1 ——   Web Server 2 ——

CPU Average Utilisation – total   11May13 – 12May13

Rules Server 1 ——   Rules Server 2 ——   Rules Server 3 ——

Rules Server 4 graph not available
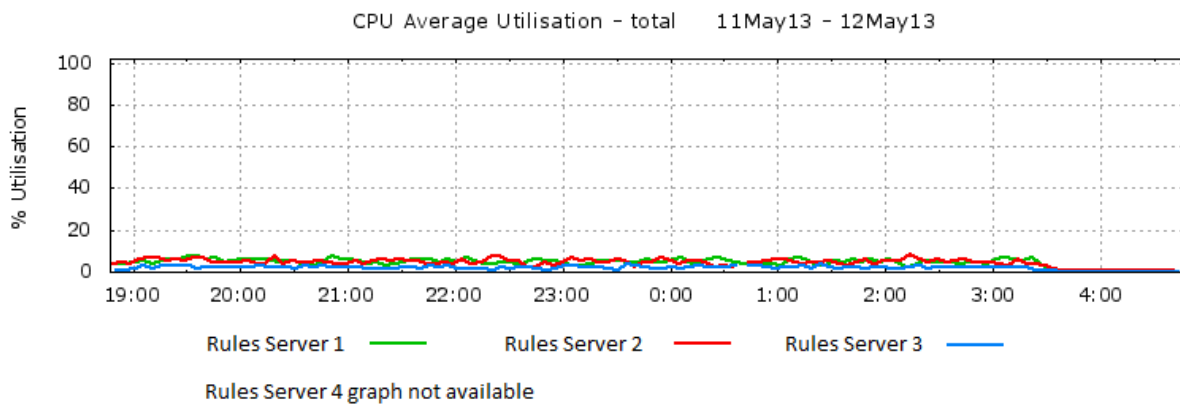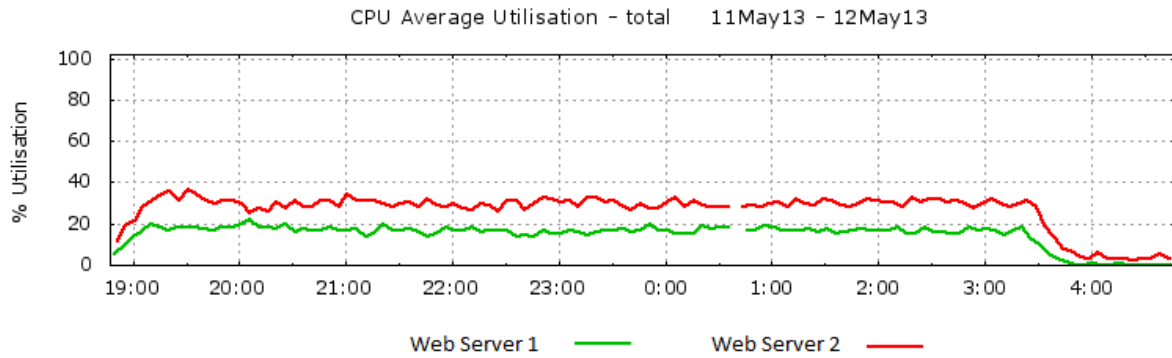
Data Server - Live Database - CPU Average Utilisation   11May13 – 12May13

total ——   user ——   priv ——

The peak at 22:00 was caused by security software updates and was not related to load testing

73

8.4. Stress Test 1

Test Parameters: 370 Concurrent Virtual users for 2 hours, target 1,090 submissions an hour.

Executed: 13/05/2013 between 17:42:37- 20:21:15

Objective: To apply more than 1.5 times the peak load on Blaise IS at different network speeds to verify if the system can sustain additional load without any issues for MPS, QBIS and REACS surveys.

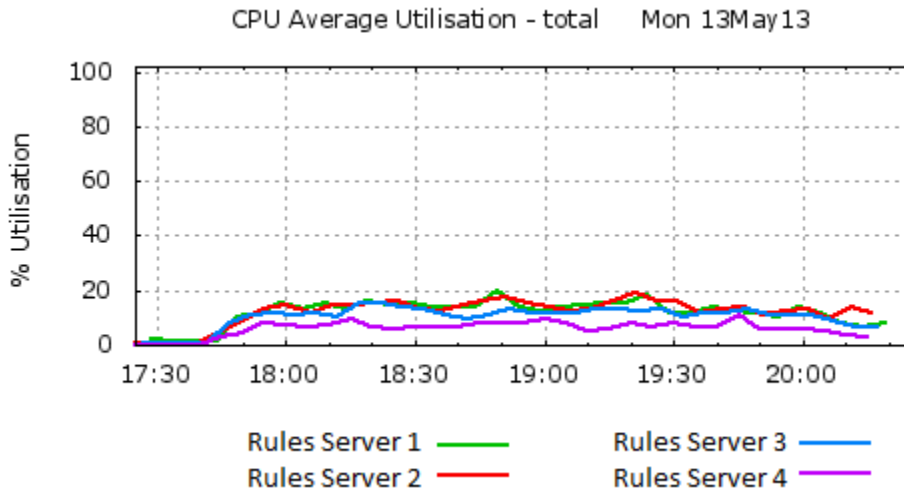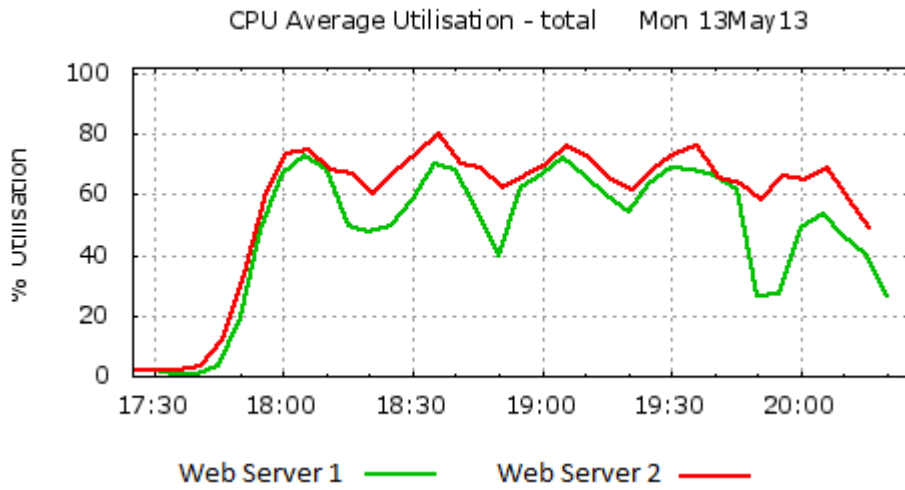The following network speeds were applied: 56Kbps, 64Kbps, 512Kbps and 2048Kbps.

Key Observations

- The total number of survey forms submitted was 940 in one hour. This test did not achieve the target survey submissions rate.

- Total surveys submitted were in one hour.

- Many errors were detected between 19:40 - 19:52.We believe this was due to connection time-outs between the Blaise API Services3 and the Journal Database.
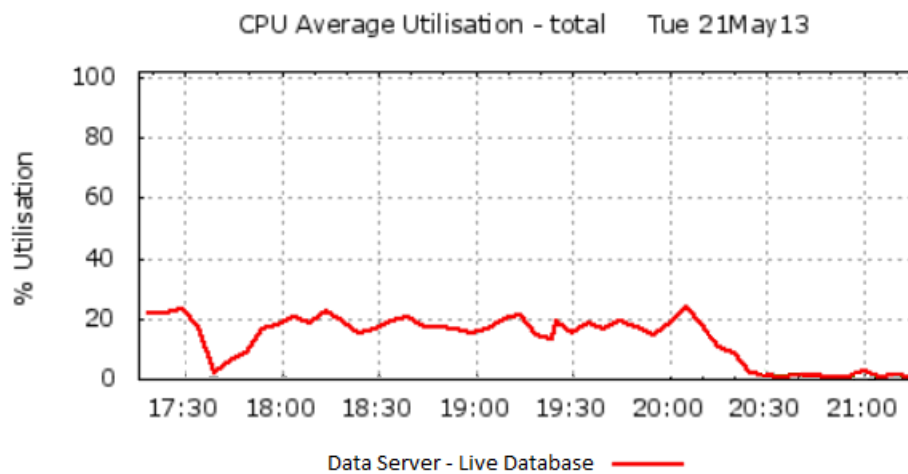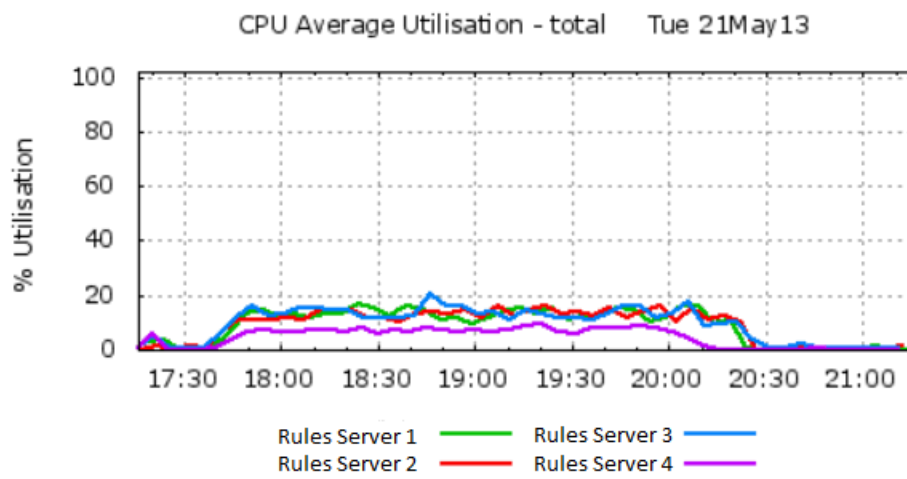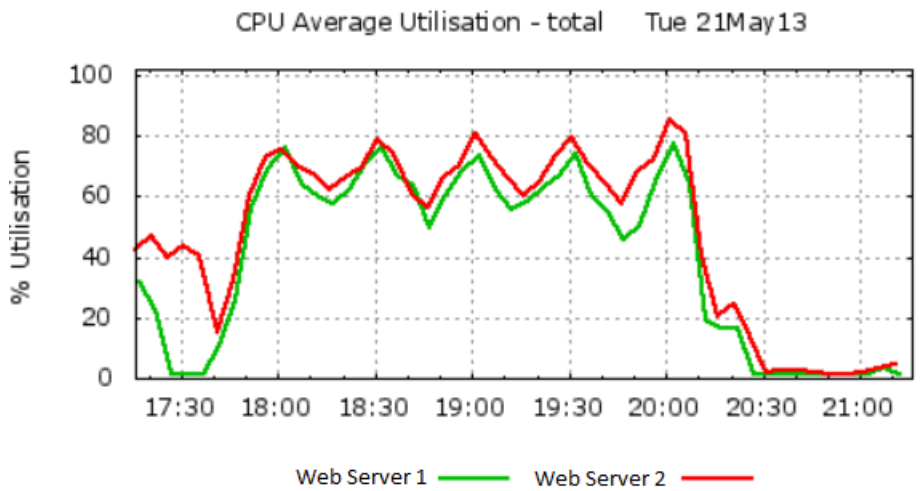
  Error: BlJour3A.Journal: Could not connect to BlaiseAPIService3 (Socket Error # 10060-Connection timed out.); ErrorNumber: -2147215301.

- 1,600 TCP/IP sockets were observed in TIME_WAIT state on the Blaise Data Server. These connection were confirmed to have originated primarily from the Blaise Web Servers

- Throughput was averaging at 1.1 Mbps.

- CPU utilization on Web Servers peaked at over 80%, and was around 20% on Rules Servers and Data Server.

The cause of the errors observed in this test was identified as the large build-up of 1,600 TCP/IP sockets in TIME_WAIT state on the Blaise Data Server. The build-up of the TCP/IP sockets was due to the number of connections that the Blaise API processes on the Web Servers were making to the Data Server for Blaise journaling calls. A fix in the form of a Windows Registry setting for the TIME_WAIT value was identified through research on the internet and applied to the Blaise Data Server (MSDN 2013; IBM 2013). The stress test was subsequently re-run successfully on 21/05/2013. Further analysis and comments are available in Section 10.5.

8.5. Stress Test 1 Graphs - Failed Attempt

CPU Average Utilisation - total   Mon 13May13

Web Server 1 ———   Web Server 2 ———

CPU Average Utilisation - total   Mon 13May13

Rules Server 1 ———   Rules Server 3 ———
Rules Server 2 ———   Rules Server 4 ———

Data Server - Live Database - CPU Average Utilisation   Mon 13May13

total ———   user ———   priv ———
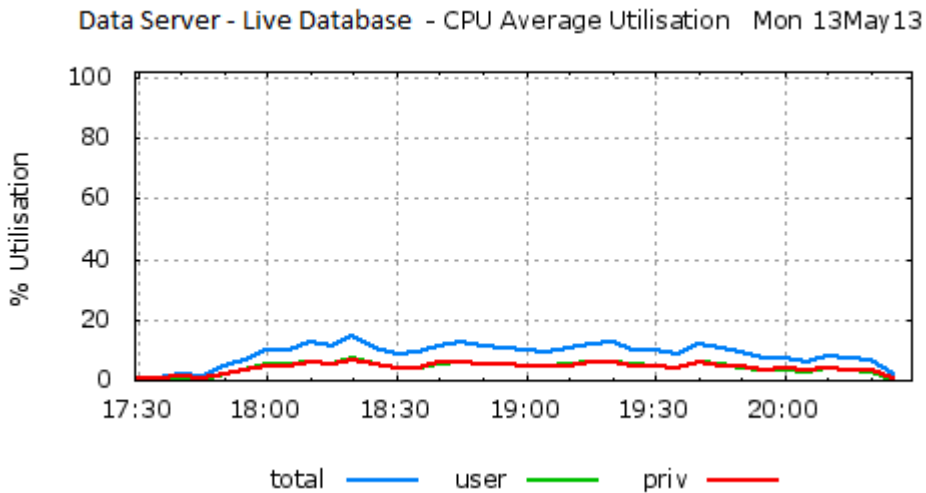
8.6. Stress Test 1 Graphs - Successful Attempt

CPU Average Utilisation - total   Tue 21May13

Web Server 1 ——   Web Server 2 ——

CPU Average Utilisation - total   Tue 21May13

Rules Server 1 ——   Rules Server 3 ——
Rules Server 2 ——   Rules Server 4 ——

CPU Average Utilisation - total   Tue 21May13

Data Server - Live Database ——

8.7. Peak (Absolute) Capacity Test

Test Parameters: 221 Concurrent Virtual users for 2 hours, target 696 submissions an hour.

Executed: 04/06/2013 between 17:19:57 - 20:25:29

Objective: Capacity (absolute maximum peak day) load was applied to Blaise eCollect system using various network speeds and benchmark end user response times for MPS, QBIS, REACS, CAPEX and ECS surveys.
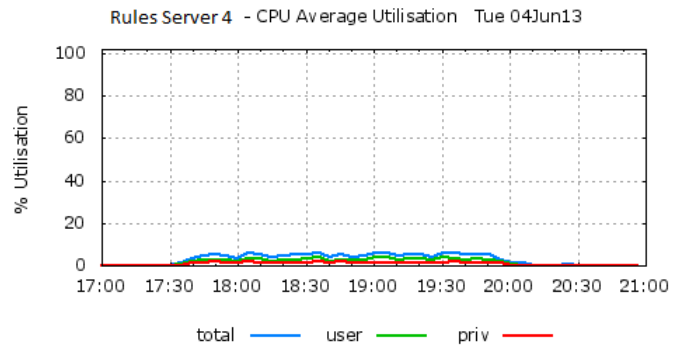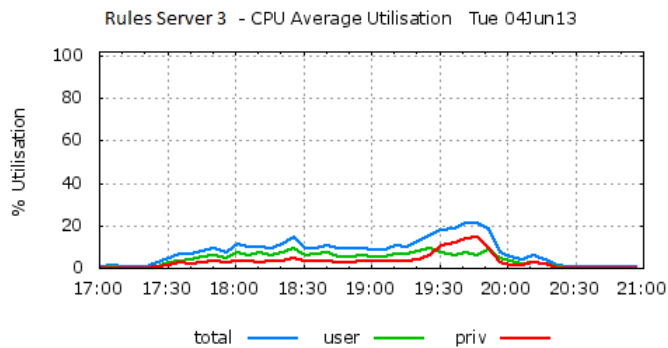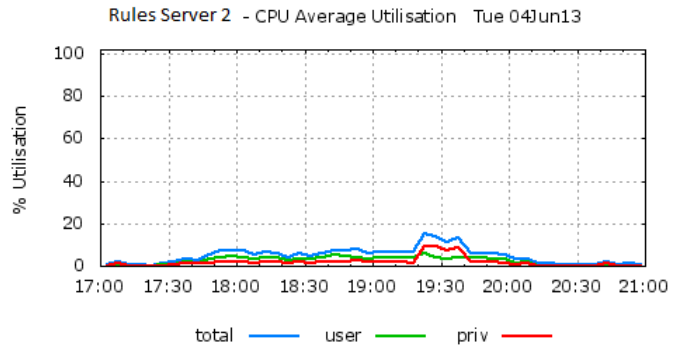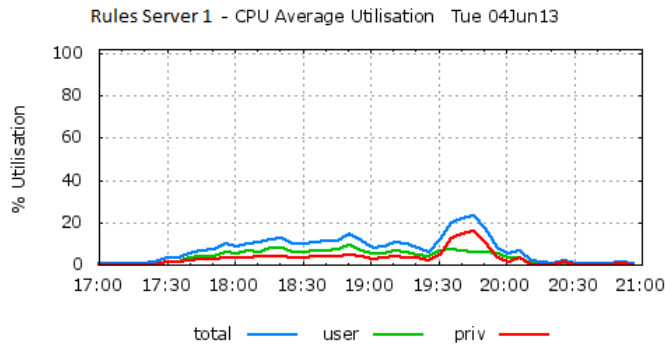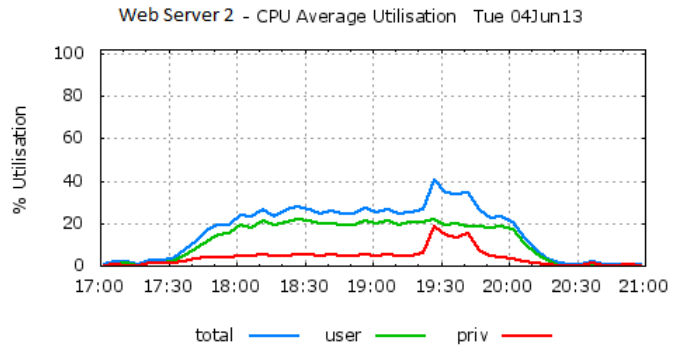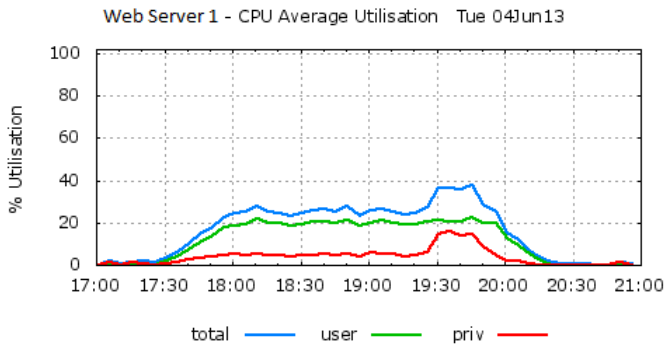
Based on results of previous stress tests the number of CPUs was increased from 2 to 4 CPUs on each of the Web Servers.

The following network speeds were applied: 56Kbps, 64Kbps, 512Kbps and 2048Kbps.

Key Observations

- The total number of survey forms submitted was 1,540 and it was as per transaction rate of 696 submissions an hour targeted for peak load test.

- There were no errors seen throughout the execution of load test run.

- Response times for 95% of the transactions for submission of survey pages at 512Kbps and 2048Kbps are within acceptable SLA of 5 secs.

- Response time for 95% Login module transactions and click survey transactions at 512Kbps and 2048Kbps is within the SLA of 15 secs.

- Response times for 95% of the transactions for submission of survey pages at 56Kbps and 64Kbps exceed SLAs of 5 seconds, and are in the range of 10-20 seconds, and as high as 40 seconds.

- Response times for 95% of the Login module transactions and click survey transactions at 56Kbps and 64Kbps exceed 15 seconds and are as high as 80 seconds.

- Throughput was averaging 4.3 Mbps.

- Total transaction throughput was averaging 6.0 transactions per second.

- CPU Utilization on web servers was averaging around 30%, and as high as 40% for a period of time.

- CPU Utilization on rules and data server was less than 20%.

8.8. Peak (Absolute) Capacity Test Graphs

Web Server 1 - CPU Average Utilisation   Tue 04Jun13

Web Server 2 - CPU Average Utilisation   Tue 04Jun13

Rules Server 1 - CPU Average Utilisation   Tue 04Jun13

Rules Server 2 - CPU Average Utilisation   Tue 04Jun13

Rules Server 3 - CPU Average Utilisation   Tue 04Jun13

Rules Server 4 - CPU Average Utilisation   Tue 04Jun13

Data Server - Live Database - CPU Average Utilisation   Tue 04Jun13

78

8.9. Stress Test 2- Excluding Authentication and Authorization

Test Parameters: 441 Concurrent Virtual users for 2 hours, target 3,097 submissions an hour.

Executed: 06/06/2013 between 21:58:03 - 00:49:31

Objective: To apply a stress test on Blaise IS using 441 concurrent virtual users and targeting 3,097 survey submissions an hour. This test was aimed at pushing the limits of the Blaise IS in its current configuration, but without the ABS authentication and authorisation module. For this test the custom ABS authentication and authorisation module was removed in order to test Blaise IS performance in its 'vanilla form' without custom modifications. The purpose of this was to see how well the platform performs without the additional load that is caused by authentication and authorization being done in Blaise.

The following network speeds were applied: 56Kbps, 64Kbps, 512Kbps and 2048Kbps.

This test was based on a single MPS survey instrument.

This test targeted a very high throughput of 3,307 submissions.

Key Observations

- Successful ramp up of 441 users.

- A lot of errors and failures were seen throughout the test run. These errors were due to out-of-memory errors reported on the Rules Servers. The target of 3,307 submissions per hour was not reached as there were many failures.

- Interestingly, while the out of memory errors were reported by the Blaise Rules Servers, the affected Rules Servers had a significant amount of available memory, at least 1GB on each Rules Server.

- The results from this test need to be investigated further. Additional comments on the results from this test are provided in Section 11 - Challenges and Issues.
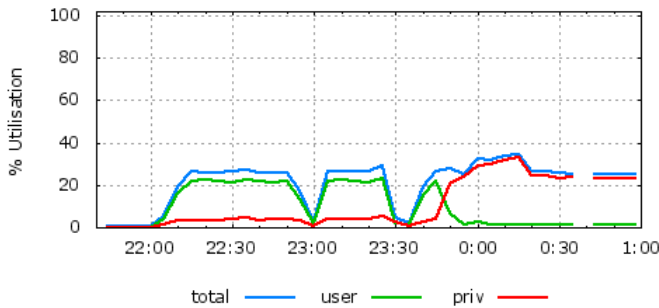
8.10. Stress Test 2 Graphs

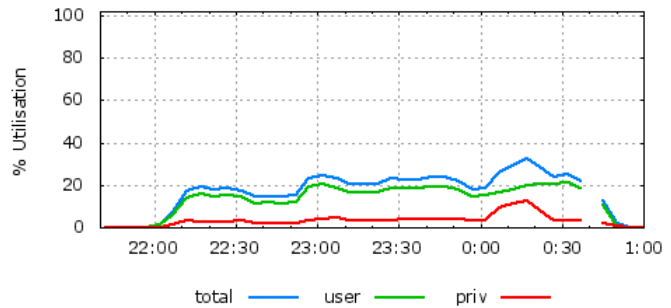Web Server 1 - CPU Average Utilisation 06Jun13 - 07Jun13
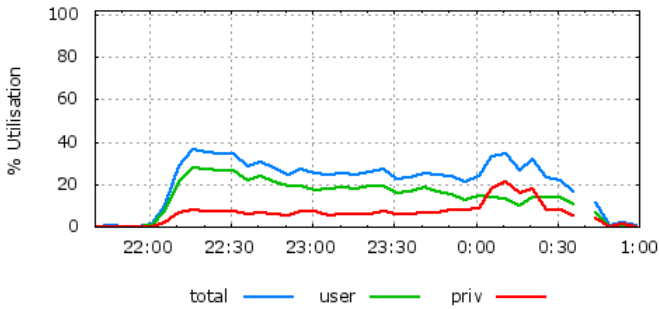
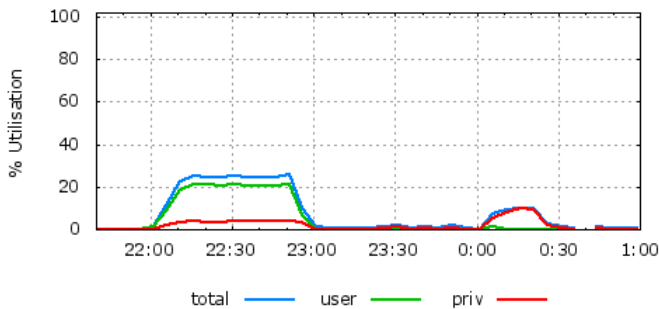Web Server 2 - CPU Average Utilisation 06Jun13 - 07Jun13
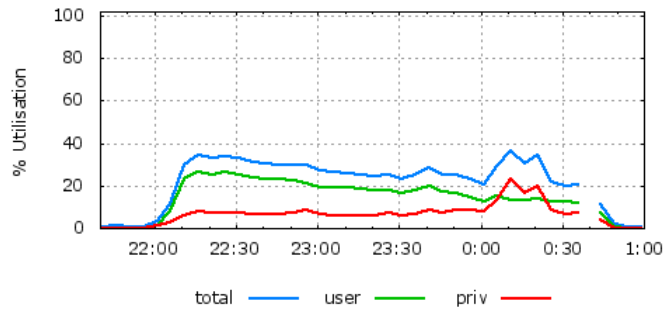
Rules Server 1 - CPU Average Utilisation 06Jun13 - 07Jun13

Rules Server 2 - CPU Average Utilisation 06Jun13 - 07Jun13

Rules Server 3 - CPU Average Utilisation 06Jun13 - 07Jun13

Rules Server 4 - CPU Average Utilisation 06Jun13 - 07Jun13

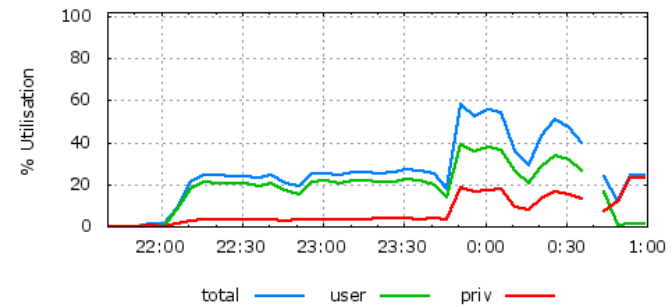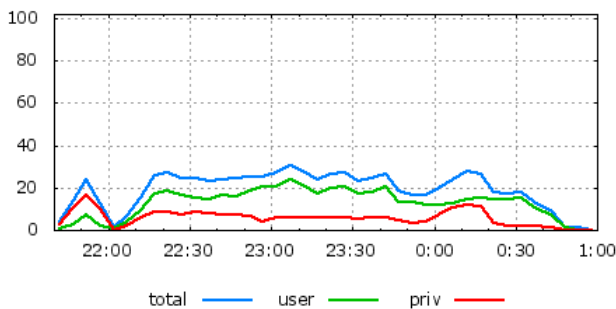Data Server - Live Database - CPU Average Utilisation 06Jun13 - 07Jun13

The gaps in if the graphs at approximately 00:45 is due to system updates when performances metrics were not being updated. It is not related to the load and performnce test.

80

8.11. Stress Test 3 - Including Authentication and Authorization

Test Parameters: 441 Concurrent Virtual users for 2 hours, target 1,397 submissions an hour

Executed: 11/06/2013 17:19:39 - 19:49:11

Objective: To apply twice the absolute peak load to Blaise eCollect system using different network speeds to verify if the system can sustain the load without any issues for MPS, QBIS, REACS, CAPEX and ECS surveys.

The following network speeds were applied: 56Kbps, 64Kbps, 512Kbps and 2048Kbps.
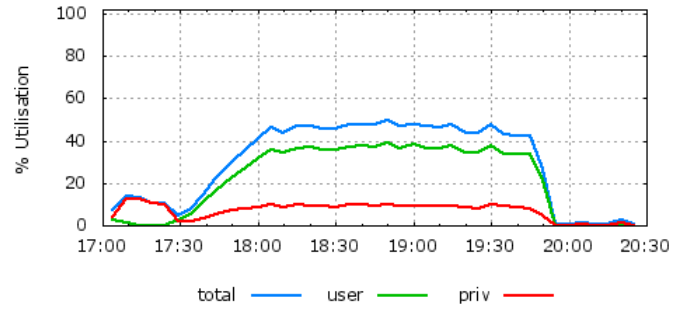
Key Observations

- Total surveys submitted were 2,795 and it was as per transaction rate of 1,397 submissions an hour targeted for peak load test.

- The total number of survey forms submitted was 3,231

- There were no errors seen throughout the execution of the test run.

- Response times for 95% of the transactions for submission of survey pages at 512Kbps and 2048Kbps are within acceptable SLA of 5 secs.

- Response time for 95% Login module transactions and click survey transactions at 512Kbps and 2048Kbps is within the SLA of 15 secs.

- Response times for 95% of the transactions for submission of survey pages at 56Kbps and 64Kbps exceed SLAs of 5 seconds, and are in the range of 10-20 seconds, and as high as 40 seconds.

- Response times for 95% of the Login module transactions and click survey transactions at 56Kbps and 64Kbps exceed 15 seconds and are as high as 80 seconds.

- Throughput was averaging at 8 Mbps.

- Total transaction throughput was averaging 13 transactions per second

- CPU Utilization on web servers was averaging around 50%.

- CPU utilization on rule severs was averaging at 30% and on data server was 20%

8.12. Stress Test 3 Graphs

Web Server 1 - CPU Average Utilisation   Tue 11Jun13
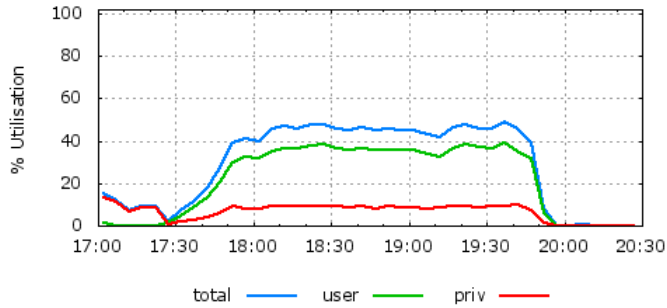
total —— user —— priv ——

Web Server 2 - CPU Average Utilisation   Tue 11Jun13

total —— user —— priv ——

Rules Server 1   CPU Average Utilisation   Tue 11Jun13

total —— user —— priv ——

Rules Server 2 - CPU Average Utilisation   Tue 11Jun13

total —— user —— priv ——

Rules Server 3 - CPU Average Utilisation   Tue 11Jun13

total —— user —— priv ——

Rules Server 4 - CPU Average Utilisation   Tue 11Jun13

total —— user —— priv ——

Data Server - Live Database - CPU Average Utilisation   Tue 11Jun13

total —— user —— priv ——

82
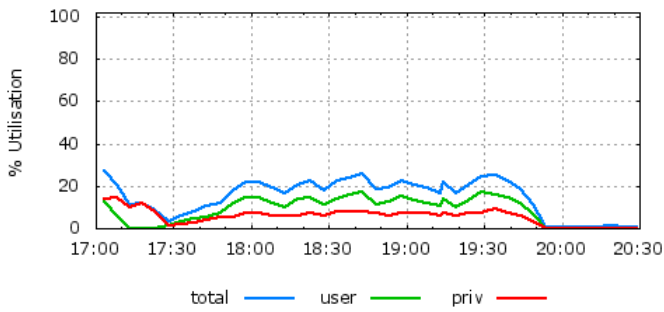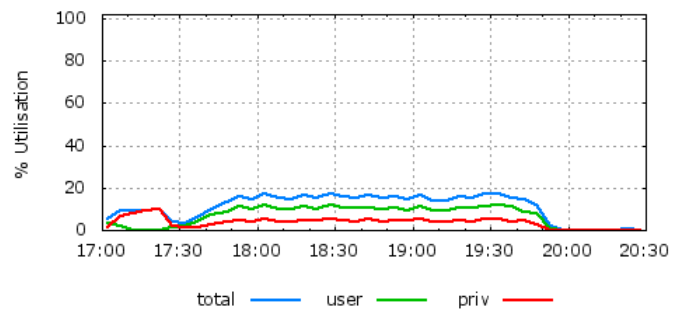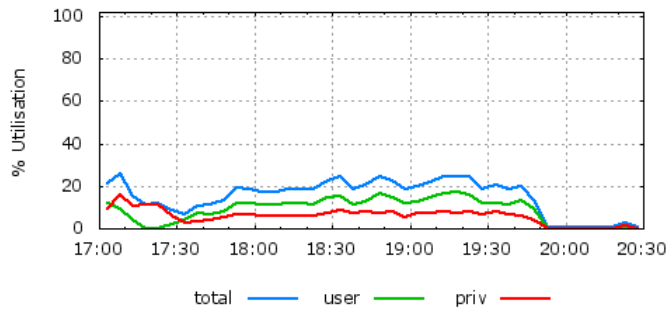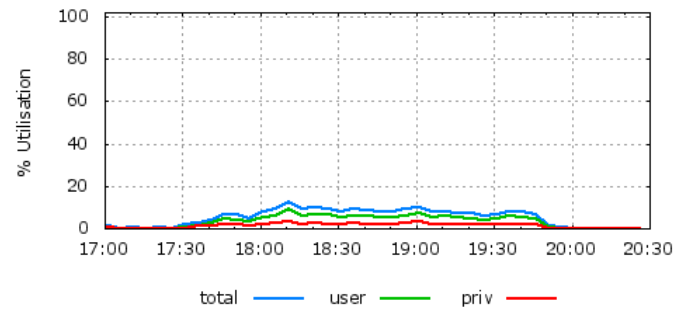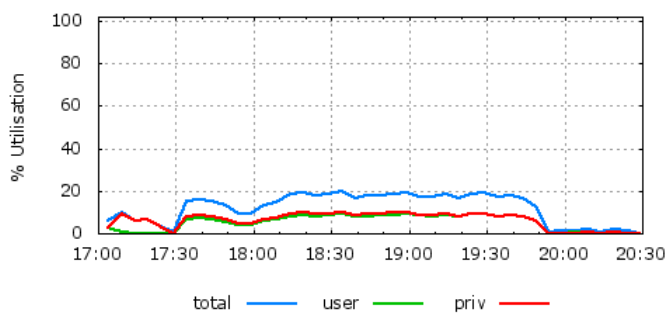
8.13. Data Extraction Test

Test Parameters: 221 Virtual users for 2 hours + Data Extraction, target 696 submissions an hour.

Executed: 03/06/2013 between 18:00:08 - 21:05:42

Objective: To apply the absolute peak load to Blaise eCollect system using various network speeds and run data extraction process in parallel to the load in order to verify the effect of data extraction on the end user response times and also to validate the performance of the data extraction module. For this test, the peak load was run for a full 1 hour before data extraction was run. This was done to create the appropriate number of records (survey submissions) for the expected peak load, and is based on the assumption that data extraction will be run every hour. It needs to be noted that data extraction is implemented as an incremental process, meaning that survey submissions that were previously extracted are not re-extracted on subsequent runs.

The following network speeds were applied: 56Kbps, 64Kbps, 512Kbps and 2048Kbps.

Key Observations

- Total surveys submitted were 1,685 and it was as per transaction rate of 696 submissions an hour targeted for peak load test.

- There were no errors seen throughout the execution of load test run.

- The data extraction module was able to handle 1 hour of data in less than 2 minutes and had negligible impact on front end system performance.

- On average it took 20 seconds to extract 300 records (survey submissions).

- Response times for 95% of the transactions for submission of survey pages at 512Kbps and 2048Kbps are within acceptable SLA of 5 secs.

- Response time for 95% Login module transactions and click survey transactions at 512Kbps and 2048Kbps is within the SLA of 15 secs.

- Response times for 95% of the transactions for submission of survey pages at 56Kbps and 64Kbps exceed SLAs of 5 seconds, and are in the range of 10-20 seconds, and as high as 40 seconds.

- Response times for 95% of the Login module transactions and click survey transactions at 56Kbps and 64Kbps exceed 15 seconds and are as high as 80 seconds.

- Throughput was averaging 4.48 Mbps.

- Total transaction throughput was averaging 4.0 transactions per second.

- CPU Utilization on web servers was averaging around 20%.

- CPU Utilization on rules and data server was less than 20%.

9. Summary of Results

The following is the overall summary of the results obtained through load and performance testing:

- Response times for at least 90% of the transactions for submission of survey pages at 512Kbps and 2048Kbps were within acceptable SLA of 5 seconds.

- Response times for at least 90% Login module transactions and click survey transactions at 512Kbps and 2048Kbps were within the SLA of 15 seconds.

- Response times for at least 90% of the transactions for submission of survey pages at 56Kbps and 64Kbps exceed SLAs of 5 seconds, and are in the range of 10-20 seconds, and as high as 40 seconds.

- Response times for at least 90% of the Login module transactions and click survey transactions at 56Kbps and 64Kbps exceed 15 seconds and are as high as 80 seconds.

- The data extraction module was able to handle 300 records (survey submissions) in under 20 seconds and had negligible impact on front end user experience.

- Blaise eCollect system was able to meet the defined twice the Peak Load for 2 hours without any performance degradation with 441 concurrent users achieving 1,397 survey submissions an hour.

- Blaise eCollect system was able to meet the defined Normal Load for prolonged period of time of 8 hours without any performance degradation such as memory leaks, user response time degradation, or any other observed negative trend with 127 users. The number of survey submissions achieved in 8 hours was 3,391.

- CPU utilization was averaging around 50% on Web Servers, 30% on Rule Severs and 20% on the Data Server for the highest load tests. On all servers memory available was sufficient throughout the test. These metrics were for the stress test at 441 concurrent users and 1,397 submissions an hour.

- In Stress Test 1, which included 370 concurrent users and a target rate of 1,090 submissions an hour, a lot of errors were observed due to the large build-up of 1,600 TCP/IP sockets in TIME_WAIT state on the Blaise Data Server. The build-up of the TCP/IP sockets was due to the number of connections that the Blaise API processes on the Web Servers were making to the Data Server for Blaise journaling calls. A fix in the form of a Windows Registry setting for the TIME_WAIT value was applied to the Data Server and the stress test was re-run successfully.

- In Stress Test 3, which included 441 concurrent users and a target rate of 3,307 survey submissions an hour a lot of and failures were seen throughout the test run. These errors were due to out-of-memory errors reported on the Rules Servers. Interestingly, while the out of memory errors were reported by the Blaise Rules Servers, the affected Rules Servers had a significant amount of available memory, at least 1GB on each Rules Server. The results from this test need

to be investigated further. Additional comments on the results from this test are provided in Section 11 - Challenges and Issues.

10. Scalability Observations and Analysis

10.1. Web Servers

In the solution architecture for Blaise eCollect, a specialised F5 load balancer was used to distribute the load to the web servers. In all the tests conducted it was observed that the load on the web servers was well distributed and evenly spread.

Also, it should be noted that the number of CPUs on each of the Web Servers was upgraded from 2 to 4 CPUs based on results of Stress Test 1 - Section 8.4, where CPU utilization on the Web Servers was observed to peak at 80%. The additional CPUs resulted in a much lower CPU utilization in further stress tests, which peaked at only 50% utilization for the same load.

Rules Servers

In the solution architecture for Blaise eCollect, the built-in Blaise IS load balancing feature was used to distribute the load between the rules servers. This feature was configured to "round-robin" load balancing setting. In all tests conducted, it was observed that the load on the rules servers was fairly well balanced and reasonably well distributed, although in some cases load distribution was not completely even.

In some of the earlier tests conducted the focus was to establish how well Blaise IS scales with additional server resources. Early on it was observed that it was the Rules Servers that took on a lot of the computational load. For example, the 2 web servers were 20% CPU utilised while the 2 rules servers were above 40% CPU utilised. Due to these observations, the number of web servers was fixed at 2 servers. The number of rules servers was varied from 1 to 3 to establish how well the performance of the rules servers scales under load. This number was increased per test setup until errors or time-outs were encountered. Based on those tests it was concluded that Blaise IS performance scales reasonably well and performance scalability appears to be linear with additional Rules Servers. The results of those tests are shown in Table 2.

Table 2 Blaise Rules Servers Scalability

Web Servers	Rules Servers	Max. Concurrent Users	Total Survey Submissions per hour
2	1	20	100
2	2	40	200
2	3	60	300

10.2. Data Server

In terms of Blaise IS performance and scalability in a single Blaise Park, the Data Server is the limiting factor. This is because, there can only be one Data Server per Blaise Park. The data server is also the

single point of failure in the solution architecture, as there are no secondary data servers that can take-over in case of primary data server failure.

In terms of scalability the limitation of a single data server can be overcome by utilising more than one Blaise Park, however this increases the overall solution complexity, as any points of integration, such as for example with an organisation's back-end processing systems will need to take into account multiple Blaise Parks. The multiple park approach will also present significant additional challenges for extremely large surveys, that is, surveys with large sample sizes, which cannot be run on single Blaise Park.

10.3. Blaise 4 - 32bit Memory Limits

As Blaise 4 runs as a 32bit Windows process there is also a practical limitation on the amount of virtual memory available to the system. In a 32bit Windows environment this is limited to 2GB (3GB under special compile time configuration). While this limitation was not encountered in any of the tests conducted to date, it is likely that the Data Server will eventually reach this limitation under a high load. It is also possible that this limit has the potential to affect the Rules and Web servers. However, this can be easily mitigated by spreading the load to additional web and rules servers.

10.4. TCP/IP Socket Build-Up on Data Server

In one of the earlier stress tests, it was observed that under a high load with 370 concurrent users targeting 1,090 submissions an hour, the users experienced a large number of time-out errors, towards the end of the 2 hour test. These errors were caused by the requests to the system that could not be completed in less than 120 seconds. The reason for this was identified as the large build-up of 1,600 TCP/IP sockets in TIME_WAIT state on the data server. The cause of this was the number of connections that the Blaise API processes on the Web Servers were making to the Data Server for Blaise journaling calls.

On further investigation it was established that the Windows 2008 R2 server on which the Data Server was hosted has a default windows registry setting of 4min for TCP/IP sockets to remain in TIME_WAIT state after the connection is closed by the client. The purpose for this is to allow any belated network packets to arrive after the connection has been closed, however some industry vendors recommend changing the default to a much lower setting of 30 seconds to improve server performance (MSDN, 2013; IBM, 2013).

The default value for the TIME_WAIT sockets was changed to 30 seconds on the data server and the stress tests were re-run successfully without any errors. The change to this setting allowed the overall throughput performance to improve significantly, up to 441 concurrent users and 1,397 submissions an hour. The number of sockets observed in TIME_WAIT state with 441 concurrent users was approximately 700.

It is however likely that this will be one of the early performance limitations in Blaise 4. Based on observations before and after the setting changes to the TIME_WAIT value, it is estimated that the limit of concurrent users could be as high as 900 and the number of hourly submissions about 3,000. However, this is only a prediction and has not been tested.

A potential longer term solution to the socket build-up issue is to implement a connection pooling mechanism, whereby connections between server components are not closed after each use but rather kept

in a connection pool and reused when needed. This technique is used in a large number of enterprise industry solutions including relational database systems (Wikipedia, 2013).

11. Challenges and Issues

11.1. Authentication and Authorisation in Blaise

The Authentication and Authorisation module in the Blaise eCollect solution presented unique challenges from the point of system performance. This module is an ABS developed component primarily used to allow Blaise instruments to interact with other non-Blaise system components. The module is implemented as a .Net C# DLL and is installed on each of the Blaise Rules Servers. This module is referenced and called through the use of Blaise 'Alien' procedure calls in Blaise instruments. The module relies on the use of the Blaise Database and DatabaseManager objects which are used to connect to and query Blaise database files.

One of the issues that was encountered with this approach to authentication and authorisation was with the concurrent calls to the module causing instability on the Blaise Rules servers and causing the Blaise API process (BlAPI3S.exe) to crash under concurrent access calls from as low as 10 simultaneous users. This issue was identified and rectified with concurrent access locking mechanisms available in .Net C#. The module has now been re-tested to support up to 500 concurrent user requests.

Given that this authentication and authorisation approach relies on calls to the Blaise database and Blaise Data Server, it places an additional load on the eCollect system. Therefore there is currently work being undertaken by ABS to utilize a dedicated authentication mechanism based outside of Blaise.

11.2. Slow performance at 56kbps and 64kbps dial-up

The transaction response times for end-user dial up connection speeds of 56kbps and 64kbps were found to be relatively high, as outlined in section.

- Response times for at least 90% of the transactions for submission of survey pages at 56Kbps and 64Kbps exceed SLAs of 5 seconds, and are in the range of 10-20 seconds, and as high as 40 seconds.

- Response times for at least 90% of the Login module transactions and click survey transactions at 56Kbps and 64Kbps exceed 15 seconds and are as high as 80 seconds.

Investigations are currently being made into improving the performance at these end-user connection speeds. This effort is focused around using the Google Closure Tools for JavaScript optimization as well as HTTP traffic compression on web servers and load balancer devices (Google, 2013).

11.3. Access Violation and Out of Memory Errors

One of the tests that was run focused on evaluating the performance of the Blaise IS solution in its 'vanilla' form. That is without the inclusion of an authentication and authorisation module developed by ABS. This test was aimed at assessing the raw performance of Blaise as a way of comparing the additional load the authentication and authorisation done through Blaise places on the system. The parameters for this test included 441 concurrent users, and survey submissions up to 3,307 an hour, over a two hour period. This test was also based on a single MPS survey instrument.

The test ran well for 30 minutes, at this point one of the Blaise Rules Servers experienced serious failures. The errors manifested as time-out errors to the end-users, where the request to process a survey operation could not be completed within 120 seconds. However, on the Blaise Rules servers the errors appeared as Windows Application errors in the Windows Application Event logs. There were several hundred of these errors reported in the logs.

The following is a short extract of the errors and the full errors list is available in the Appendix.

Error 1: TBIAPIManager.ParseXMLDoc. E.Message: Unrecognized exception: Access violation at address 00401E6F in module 'BIAPI3S.exe'. Read of address 00000000 E.ErrorCode: -2147192832 E.ErrorSource: Database: 898817088 Catastrophic: true

Error 2: 10:24:44.462 TBIAPIManager.ParseXMLDoc. E.Message: Unrecognized exception: Out of memory E.ErrorCode: -2147192832 E.ErrorSource: Database: 373684400 Catastrophic: false

This test was repeated two times with the same overall results each time. On the second run through, two of the four Blaise Rules servers experienced these errors and the simulated users failed to complete the test scenarios.

Interestingly, while the out of memory error was reported by the Blaise application, the affected Rules Servers had a significant amount of available memory, at least 1GB.

The results from this test need to be investigated further. There were a number of key parameters in this test that need further exploration. Specifically, this test had a higher target of 3,307 survey submissions an hour which could be a possible explanation for the observed errors. In addition, this test used a single MPS survey instrument, where previously, multiple surveys were used, and MPS users represented a much smaller proportion of all survey respondents. The MPS has a large and complex hierarchical question structure which can potentially have an impact on memory utilization with a large number of survey respondents.

The findings from this test were shared with the Blaise team from Statistics Netherlands. Currently, there is no confirmed explanation for these errors. However, the Statistics Netherlands team and ABS are investigating this issue further.

12. Conclusion

In conclusion, the Blaise 4.8.4 eCollect platform performs well at 512kbps and 2048kpbs end-user network speeds, meeting the performance targets of 5 seconds for survey navigation transactions and 15seconds for login transactions. The platform performs well for concurrent users, supporting 441 concurrent users for 2 hours and achieving over 1,397 submissions an hour. The system also performed well in endurance tests, achieving a throughput of 3,307 submissions in 8 hours with 127 concurrent users and maintaining responsive transaction times and stable performance characteristics.

Challenges still remain in improving system performance for dial-up connections where transaction times were reported at up to 40 seconds for survey navigation transactions and 80 seconds for login transactions. In addition, the out-of-memory errors observed at a higher level of survey submission throughput are an issue and will require further investigation and resolution. Finally, the scalability of a Blaise Park is challenged by the single Data Server, which is further compounded by issues seen with end-user request time-outs caused by TCP/IP socket build-up from journaling API calls between Blaise Web Servers and the Data Server at higher levels of system load. Whilst additional Blaise Parks may allow further scalability, this may present additional complexities for surveys with very large numbers of respondents.

Challenges and issues encountered during the test program will need to be considered and addressed, particularly in the areas of the out of memory reports seen, and poor performance using slower connection speeds.

Overall, the load and performance test program allowed the ABS to gain valuable insights into performance, capacity and scalability of the Blaise 4.8.4 eCollect platform. Based on the test results it was concluded that the eCollect platform will have sufficient resources and scalability to support the ABS June 2013 quarter eForm survey migration goals. Furthermore, there is also potential for further scalability to support additional ABS eForm surveys into the future.

13. References

1. Google, Google Closure JavaScript Optimiser, Accessed: 22/07/2013,
    https://developers.google.com/closure/

2. MSDN, Avoiding TCP/IP Port Exhaustion, Accessed: 22/07/2013, http://msdn.microsoft.com/en-us/library/aa560610(v=bts.20).aspx

3. IBM, Configuring Windows for high network connection rates, Accessed: 22/07/2013,
    http://publib.boulder.ibm.com/infocenter/cicstg/v6r0m0/index.jsp?topic=%2Fcom.ibm.cicstg600.doc%2Fccllal0264.htm

4. Wikipedia, Connection Pool, Accessed: 23/07/2013 http://en.wikipedia.org/wiki/Connection_pool

5. ABS, Internet Activity Survey Dec 2012, Accessed: 22/07/2013,
    http://www.abs.gov.au/ausstats/abs@.nsf/mf/8153.0/

14. Appendix

14.1. Surveys Selected for Testing

Table 3 Surveys Selected for Testing

Survey	Expected eForm Take-up	Number of Question Fields	Reason to include in testing
MPS (Monthly Population Survey)	11,000	400	Representative of population/household survey structure. Monthly cycle applies more frequent load on the system
CAPEX (Capital Expenditure Survey)	8,648	26	Quarterly survey, chosen as representative of business survey structure, with large sample size.
QBIS (Quarterly Business Indicators Survey)	16,691	50	Quarterly survey, chosen as representative of business survey structure, with large sample size.
ECS (Engineering Construction Survey)	4,409	140	Quarterly survey, chosen as representative of business survey structure, with large sample size.
REACS (Rural Environment and Agricultural Commodities Survey)	35,000	470	Annual survey. Chosen for its length (number of questions) and large sample size

14.2. Internet Connection Speeds

Table 4* Internet Connection Speeds

Speed	No. of Connections	Percentile %
56Kb-dial-up-modem	282,000	2.32
64Kb-ISDN	6000	0.5
512Kb-Satellite (900ms latency)**	92,000	0.76
512Kb-ADSL	4,727,000	38.42
2048Kb-Other	7,060,000	58
Total	12,167,000	100

*This table was based on the ABS Internet Activity Survey Dec 2012. http://www.abs.gov.au/ausstats/abs@.nsf/mf/8153.0/

**Note: 512Kb- Satellite (900ms latency) could not be simulated via LoadRunner 11 for load and performance tests.

14.3. Out of Memory Error Logs

Windows Event Application Log 1

Log Name: Application
Source: BlaiseAPIService3
Date: 05/06/13 20:18:57
Event ID: 1001
Task Category: None
Level: Error
Keywords: Classic
User: N/A
Computer: rulesserver
Description:
10:18:57.587 TBlAPIManager.ParseXMLDoc. E.Message: Unrecognized exception: Access violation at address 00401E6F in module 'BlAPI3S.exe'. Read of address 00000000 E.ErrorCode: -2147192832 E.ErrorSource: Database: 898817088 Catastrophic: true
Event Xml:
<Event xmlns="http://schemas.microsoft.com/win/2004/08/events/event">
  <System>
   <Provider Name="BlaiseAPIService3" />
   <EventID Qualifiers="0">1001</EventID>
   <Task>0</Task>
   <Keywords>0x80000000000000</Keywords>
   <TimeCreated SystemTime="2013-06-05T10:18:57.000000000Z" />
   <EventRecordID>13830</EventRecordID>
   <Channel>Application</Channel>
   <Computer>rulesserver</Computer>
   <Security />
  </System>
  <EventData>
   <Data>10:18:57.587 TBlAPIManager.ParseXMLDoc. E.Message: Unrecognized exception: Access violation at address 00401E6F in module 'BlAPI3S.exe'. Read of address 00000000 E.ErrorCode: -2147192832 E.ErrorSource: Database: 898817088 Catastrophic: true</Data>
  </EventData>
</Event>

Windows Event Application Log 2

Log Name: Application
Source: BlaiseAPIService3
Date: 05/06/13 20:19:02
Event ID: 1001
Task Category: None
Level: Error
Keywords: Classic
User: N/A
Computer: rulesserver
Description:
SaveToStream: Access violation at address 00401E6F in module 'BlAPI3S.exe'. Read of address 00000000
Event Xml:
<Event xmlns="http://schemas.microsoft.com/win/2004/08/events/event">
 <System>
  <Provider Name="BlaiseAPIService3" />
  <EventID Qualifiers="0">1001</EventID>
  <Task>0</Task>
  <Keywords>0x80000000000000</Keywords>
  <TimeCreated SystemTime="2013-06-05T10:19:02.000000000Z" />
  <EventRecordID>13831</EventRecordID>
  <Channel>Application</Channel>
  <Computer>rulesserver</Computer>
  <Security />
 </System>
 <EventData>
  <Data>SaveToStream: Access violation at address 00401E6F in module 'BlAPI3S.exe'. Read of address 00000000</Data>
 </EventData>
</Event>

Windows Event Application Log 3

Log Name: Application
Source: BlaiseAPIService3
Date: 05/06/13 20:24:44
Event ID: 1001
Task Category: None
Level: Error
Keywords: Classic
User: N/A
Computer: rulesserver
Description:
10:24:44.462 TBlAPIManager.ParseXMLDoc. E.Message: Unrecognized exception: Out of memory E.ErrorCode: -
2147192832 E.ErrorSource: Database: 373684400 Catastrophic: false
Event Xml:

<Event xmlns="http://schemas.microsoft.com/win/2004/08/events/event">
  <System>
   <Provider Name="BlaiseAPIService3" />
   <EventID Qualifiers="0">1001</EventID>
   <Task>0</Task>
   <Keywords>0x80000000000000</Keywords>
   <TimeCreated SystemTime="2013-06-05T10:24:44.000000000Z" />
   <EventRecordID>13855</EventRecordID>
   <Channel>Application</Channel>
   <Computer>rulesserver</Computer>
   <Security />
  </System>
  <EventData>
   <Data>10:24:44.462 TBlAPIManager.ParseXMLDoc. E.Message: Unrecognized exception: Out of memory
E.ErrorCode: -2147192832 E.ErrorSource: Database: 373684400 Catastrophic: false</Data>
  </EventData>
</Event>

Windows Event Application Log 4

Log Name: Application
Source: BlaiseAPIService3
Date: 05/06/13 20:24:45
Event ID: 1001
Task Category: None
Level: Warning
Keywords: Classic
User: N/A
Computer: rulesserver
Description:
IdTCPServerExecute: Out of memory
Event Xml:
<Event xmlns="http://schemas.microsoft.com/win/2004/08/events/event">
  <System>
    <Provider Name="BlaiseAPIService3" />
    <EventID Qualifiers="0">1001</EventID>
    <Task>0</Task>
    <Keywords>0x80000000000000</Keywords>
    <TimeCreated SystemTime="2013-06-05T10:24:45.000000000Z" />
    <EventRecordID>13895</EventRecordID>
    <Channel>Application</Channel>
    <Computer>rulesserver</Computer>
    <Security />
  </System>
  <EventData>
    <Data>IdTCPServerExecute: Out of memory</Data>
  </EventData>
</Event>