QlikView

QLIKVIEW SCALABILITY BENCHMARK WHITE PAPER

Measuring Business Intelligence Throughput on a Single Server

QlikView Scalability Center Technical White Paper

December 2012

qlikview.com

QLIKVIEW THROUGHPUT ON A SINGLE MEDIUM SIZED SERVER

1,250 concurrent users asking and answering their own streams of questions

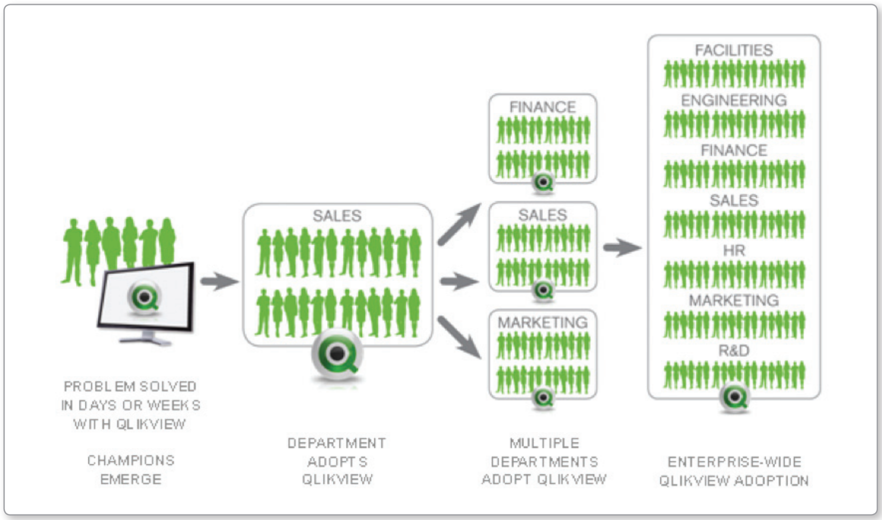
122,578 selections generated equivalent to thousands of business questions

The average response time **less than 0.4 seconds**

Executive Summary

One of the major QlikView differentiators is QlikView customers' Business Discovery adoption path. QlikView penetrates enterprises by solving significant business problems that traditional BI or data visualization tools can't address and the penetration usually starts with personal edition. In just days or weeks, QlikView solves a workgroup's business problem, perhaps a problem that would have taken months or years to solve with traditional BI, in months more QlikView apps emerge, and QlikView starts spreading across departments becoming the enterprise business intelligence solution (Figure 1). It's important to note that IT maintains control over critical elements such as data governance, data access, user rights and so on during this expansion.

Figure 1. The Typical QlikView Customer Adoption Path


PROBLEM SOLVED IN DAYS OR WEEKS WITH QLIKVIEW
CHAMPIONS EMERGE
SALES
DEPARTMENT ADOPTS QLIKVIEW
FINANCE
SALES
MARKETING
MULTIPLE DEPARTMENTS ADOPT QLIKVIEW
FACILITIES
ENGINEERING
FINANCE
SALES
HR
MARKETING
R&D
ENTERPRISE-WIDE QLIKVIEW ADOPTION

Today there are thousands of QlikView enterprise customers who have gone through this adoption path. The Business Discovery platform at these customers supports thousands of users, with very large data size and thousands of QlikView apps. One of the key ingredients to achieve this successful penetration is QlikView scalability.

In this paper, we are going to walk you through two performance test scenarios proving the business intelligence throughput that can be achieved with QlikView on a **single medium size server**. During the tests, we demonstrated how QlikView Server would support **1,250 concurrent** users asking and answering their own streams of questions. In 45 minutes, they generated over **122,578 selections** which would be equivalent to thousands of business questions only by using a single server. The average response was **less than 0.4 seconds**.

Results for a small QlikView app:

Figure 2. Scalability Test Results for a Simple QlikView App

Test Scenario Details	Results
Number of server	1
Cost of server	$16,289*
Number of sessions	4,657
Number of concurrent sessions	1,250
Average response time	0.4 seconds
Number of selections generated	122,578

* The list price for the IBM x3650 server with 192 GB of RAM and 16 2.60GHz cores as of October 26, 2012

Results for a medium QlikView app:

Figure 3. Scalability Test Results for a Medium QlikView App

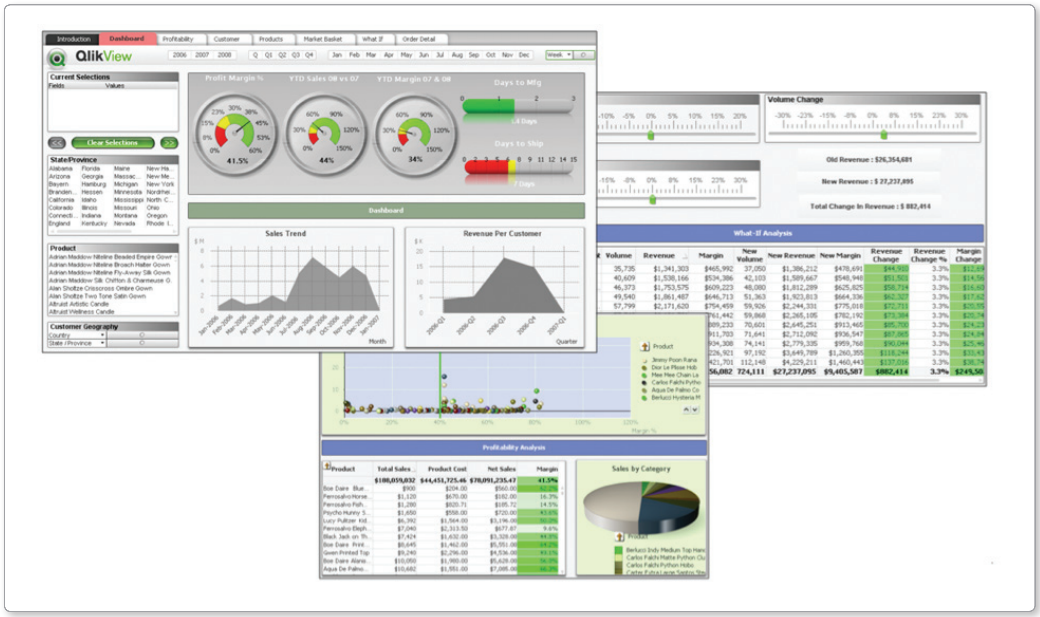
Test Scenario Details	Results
Number of server	1
Cost of server	$16,289*
Number of sessions	2,139
Number of concurrent sessions	515
Average response time	1.4 seconds
Number of selections generated	35,896

* The list price for the IBM x3650 server with 192 GB of RAM and 16 2.60GHz cores as of October 26, 2012

Please note that QlikView provides clustering capability to leverage the power of multiple servers but our goal in this paper is to prove how much more business intelligence outputs QlikView can deliver even with a single medium size server compared to traditional business intelligence or data visualization tools.

During the tests, we simulated real-world business discovery usage by using two QlikView apps, one small and one medium level app. Both of these QlikView apps had seven analysis tabs, including dashboard, profitability, products, market basket, what if and order details views. Each tab had several charts (Figure 2).

Figure 4. QlikView Apps Used During Scalability Tests



It is important to notice that the tests conducted did not just use a single tab or a view with a couple of visualizations. We used real-world production QlikView apps, with multiple tabs and dozens of charts, simulating users asking and answering streams of questions. The simulation program ran activities such as applying selections, lassoing data in charts, opening tabs, manipulating sliders, applying what-if analysis, utilizing different aggregations on charts, basically any business discovery activities that a regular business user would perform.

We believe that the scalability of a business intelligence tool should be measured in regards to the business intelligence output that it can achieve in addition to the data size and number of users supported. Traditionally, you may have seen scalability results from other BI or data visualization tools, where the tests have been conducted by using a couple of reports or a single workbook with only a couple of visualizations. But what happens in real word when the users get a hold of a traditional report or standalone data visualization and start to look at the numbers? They think of questions. No matter what their original question was, seeing the data always brings more questions to mind. And when this happens, they open another report or visualization workbook, which are excluded from the traditional scalability tests. That is why we believe traditional scalability tests do not provide an accurate performance measurement from start to end, reflecting real world analysis activities.

Therefore for us, scalability is not only a matter of the data size or number of users. It is also the business intelligence throughput that can be achieved on a given hardware where the users can gain 360 degree view on a business problem. With the testing summarized on this paper, we demonstrated how 1,250 concurrent users were able to ask and answer their own streams of questions. In 45 minutes, they generated over 122,578 selections which would be equivalent to thousands of business questions only by using a single server and getting answers in less than 0.4 seconds on average.

The next sections of the paper provide more details on the scalability testing conducted to prove QlikView's pragmatic approach to scalability.

Although these tests show results from a test environment, they reflect performance of real world production QlikView apps. The simulation software used during the tests, conducted business discovery activities that a regular user would do and simulated thousand of selections during the test time frame. However, please note that actual results may vary based on a number of variables including, load type, calculation complexity, hardware, and network speed.

Scalability Test Details

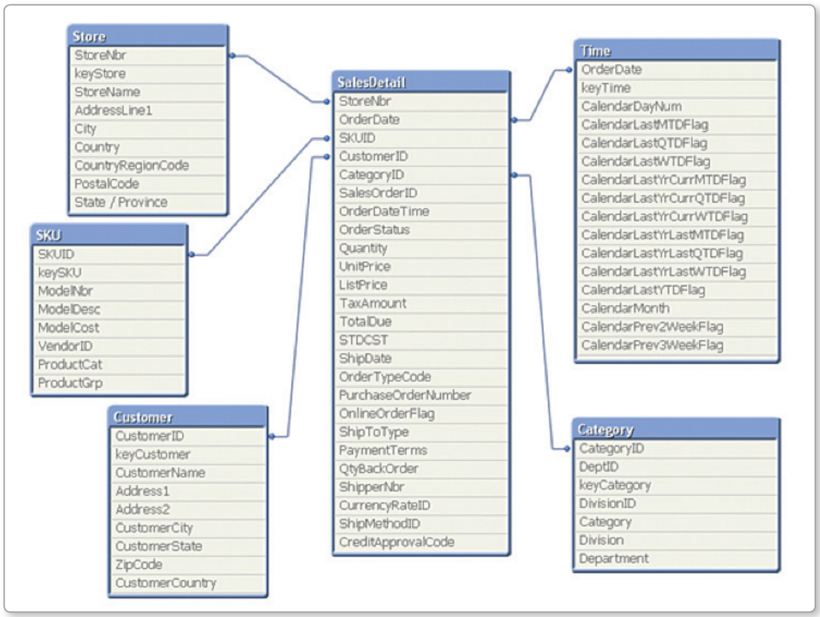
The tests that are explained on this paper are conducted at QlikTech's Scalability Labs located in Lund, Sweden. QlikTech Scalability Center is dedicated to work on topics related to performance and scalability of QlikView and provide strategic guidance to R&D on this subject. For the set of tests explained on this paper, our enterprise architects worked with the scalability labs to create real world business discovery QlikView apps.

Our testing variables were the number of user sessions, the number of concurrent users, user types, the data sizes and the complexity of the QlikView apps. For our tests, we used the same medium size server, and ran two different scenarios. As you would notice at the test results section, the system did not saturate with either of the tests and QlikView Server was able to perform very well by only using some of the available hardware resources, including RAM and CPU.

DATA

For these scalability tests, we used a real world data model supporting sales analysis. Figure 6 displays the data model. It has six tables with one fact table and 5 dimension tables. The fact table had 25 columns. The small test scenario had 10 million rows on the fact table. The medium scenario had 50 million rows. The data had very high cardinality (>95% unique data), making QlikView apps larger and more complex. The data set used during these tests can be made available upon requests.

Figure 5. QlikView App Data Model



Store
StoreNbr
keyStore
StoreName
AddressLine1
City
Country
CountryRegionCode
PostalCode
State / Province

SKU
SKUID
keySKU
ModelNbr
ModelDesc
ModelCost
VendorID
ProductCat
ProductGrp

Customer
CustomerID
keyCustomer
CustomerName
Address1
Address2
CustomerCity
CustomerState
ZipCode
CustomerCountry

SalesDetail
StoreNbr
OrderDate
SKUID
CustomerID
CategoryID
SalesOrderID
OrderDateTime
OrderStatus
Quantity
UnitPrice
ListPrice
TaxAmount
TotalDue
STDCST
ShipDate
OrderTypeCode
PurchaseOrderNumber
OnlineOrderFlag
ShipToType
PaymentTerms
QtyBackOrder
ShipperNbr
CurrencyRateID
ShipMethodID
CreditApprovalCode

Time
OrderDate
keyTime
CalendarDayNum
CalendarLastMTDFlag
CalendarLastQTDFlag
CalendarLastWTDFlag
CalendarLastYrCurrMTDFlag
CalendarLastYrCurrQTDFlag
CalendarLastYrCurrWTDFlag
CalendarLastYrLastMTDFlag
CalendarLastYrLastQTDFlag
CalendarLastYrLastWTDFlag
CalendarLastYTDFlag
CalendarMonth
CalendarPrev2WeekFlag
CalendarPrev3WeekFlag

Category
CategoryID
DeptID
keyCategory
DivisionID
Category
Division
Department

During the tests, as the simulation program generated user selections. QlikView Server was instantly calculating the aggregations for the different selected dimension combinations on average in less than 0.4 seconds for the small test and 1.4 seconds for the medium test. The results were displayed on the charts by different break downs with the associative data representation.

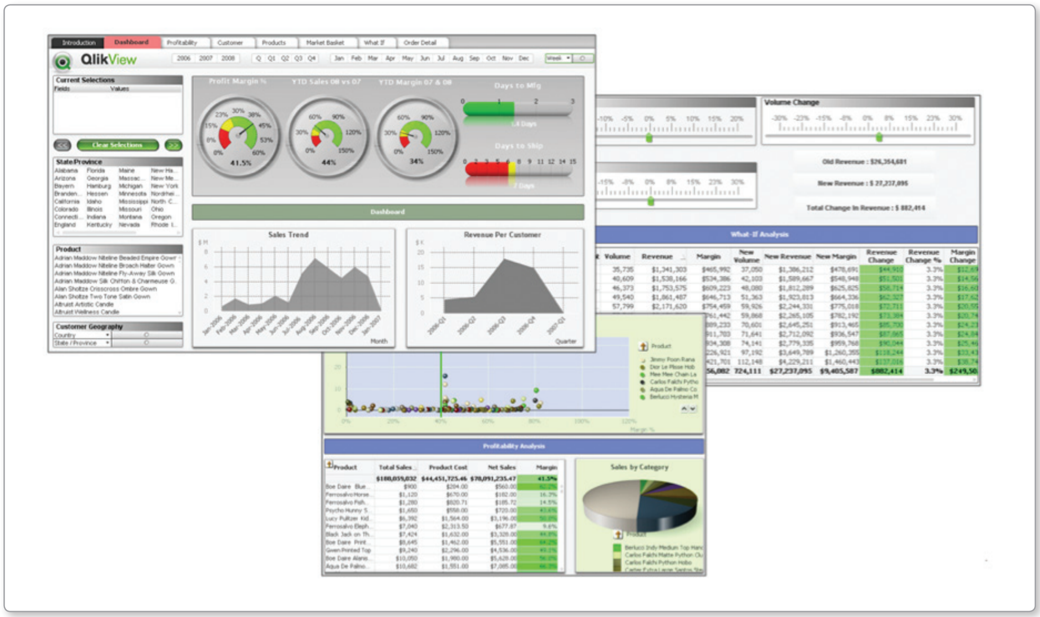
Figure 6. Data Size and Response Time Details

Results	Number of Rows on the Fact Table	Number of Sessions	Number of Concurrent Sessions	Number of Selections Simulated	Average Response Time
Small QlikView App	10 million	4,647	1,250	122,578	0.4 seconds
Medium QlikView App	50 million	2,139	515	35,896	1.4 seconds

QLIKVIEW APPS

Most organizations have a multitude of QlikView apps with varying degrees of complexity. We wanted to reflect real world use scenarios during the tests. We used two QlikView apps with different number of rows with the same analysis user interface to keep the analytics complexity the same. Each QlikView app had seven tabs including dashboard, profitability, products, market basket, what if and order details views (Figure 7).

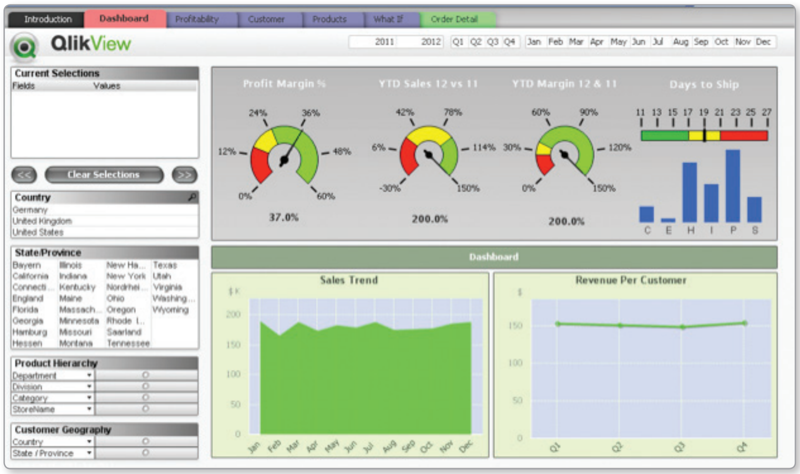
Figure 7. QlikView App Details Used During Scalability Tests.



Each tab had several QlikView charts with different visualizations and different dimensionality. As you can see from the tab descriptions, each QlikView app used during the tests can actually replace tens of traditional BI reports or data visualizations views. We conducted the performance tests with these QlikView apps to demonstrate the BI throughput that can be achieved with QlikView. This is a pragmatic approach for scalability testing compared to the other BI scalability tests that you might have seen where the vendors just use a couple of reports or visualizations to prove their scalability. The analysis details of each tab are provided below;

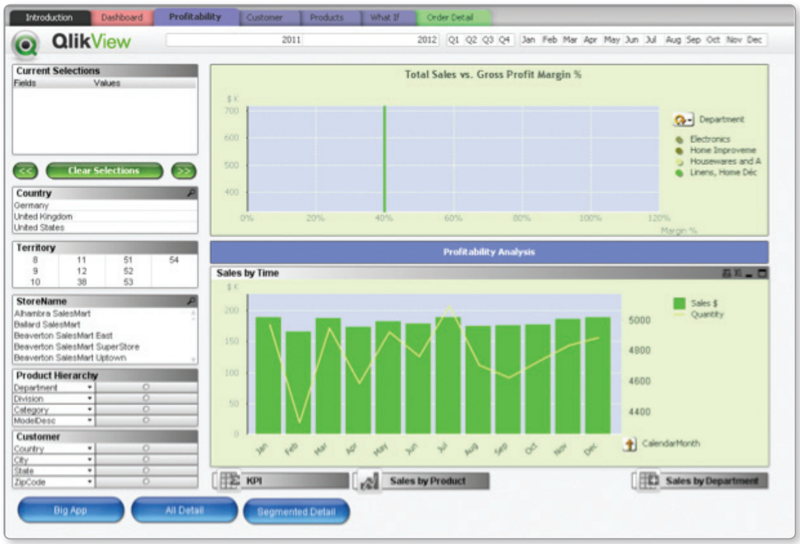
- **Dashboard tab:** It provided analysis on profit margin %, YTD sales comparison, YTD margin comparison, days to ship metric by order status, sales trend and revenue per customer. The selections available on the tab were; country, state, department, division, category, store name, customer geography and time selections with year, quarter, and month fields (Figure 8).

Figure 8. Dashboard tab of the QlikView Apps Used During Scalability Tests



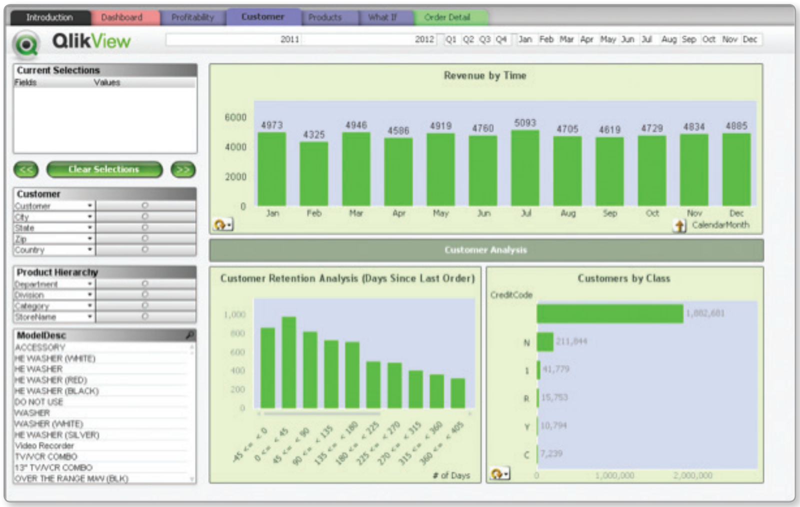
- **Profitability tab:** It provided analysis on total sales vs. gross profit margin %, sales by time, sales $, sales % to total, margin $, margin % to total and accumulated sales. The selections available on the tab were; country, territory, store name, model description, city, state, zip code, department, division, and category fields (Figure 9).

Figure 9. Profitability tab of the QlikView Apps Used During Scalability Tests



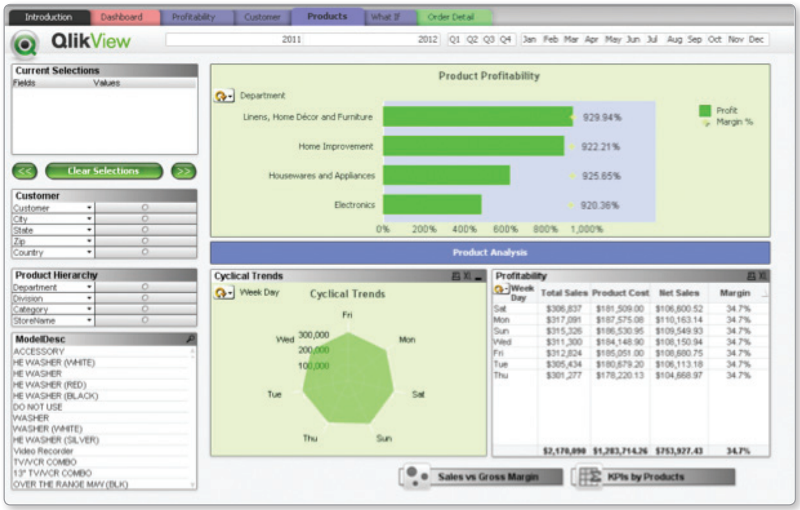
- **Customer tab:** This tab provided analysis for revenue by time, customer retention, and the number of customers by class. The selections available on the tab were; country, territory, store name, model description, city, state, zip code, department, division, category, year, quarter, and month fields (Figure 10).

Figure 10. Customer Tab of the QlikView Apps Used During Scalability Tests



- **Product tab:** It provided analysis on product profitability, cyclical trends, total sales, product cost, net sales, gross profit % and margin by year, month, week and day. The selections available on the tab were; country, territory, store name, model description, city, state, zip code, department, division, category fields (Figure 11).

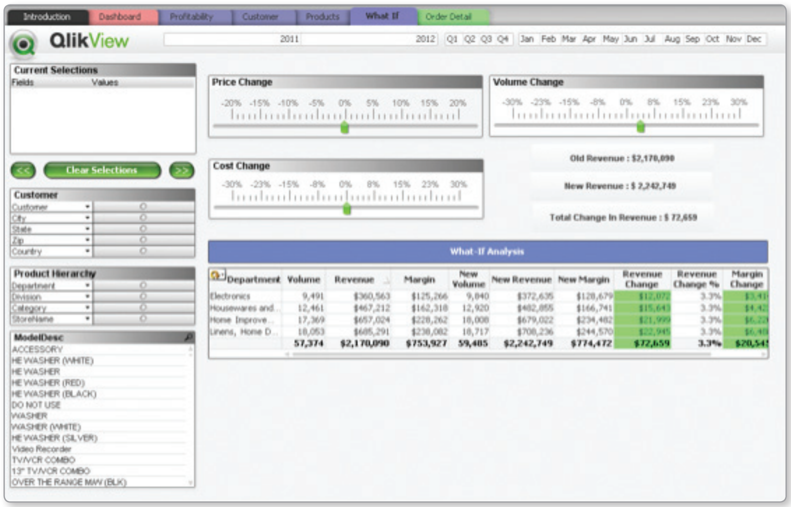
Figure 11. Product tab of the QlikView Apps Used During Scalability Tests



Week Day	Total Sales	Product Cost	Net Sales	Margin
Sat	$306,837	$181,509.00	$106,600.52	34.7%
Mon	$317,091	$187,575.08	$110,163.14	34.7%
Sun	$315,326	$186,530.95	$109,549.93	34.7%
Wed	$311,300	$184,148.90	$108,150.94	34.7%
Fri	$312,824	$185,051.00	$108,680.75	34.7%
Tue	$305,434	$180,679.20	$106,113.18	34.7%
Thu	$301,277	$178,220.13	$104,668.97	34.7%
	$2,170,090	$1,283,714.26	$753,927.43	34.7%

- **What-If analysis tab:** This tab provided what-if analysis on margin, and revenue by simulating changes on price, cost and volume. The selections available on the tab were; country, territory, store name, model description, city, state, zip code, department, division, category, year, quarter, and month fields (Figure 12).

Figure 12. What-If Tab of the QlikView Apps Used During Scalability Tests



Old Revenue : $2,170,090

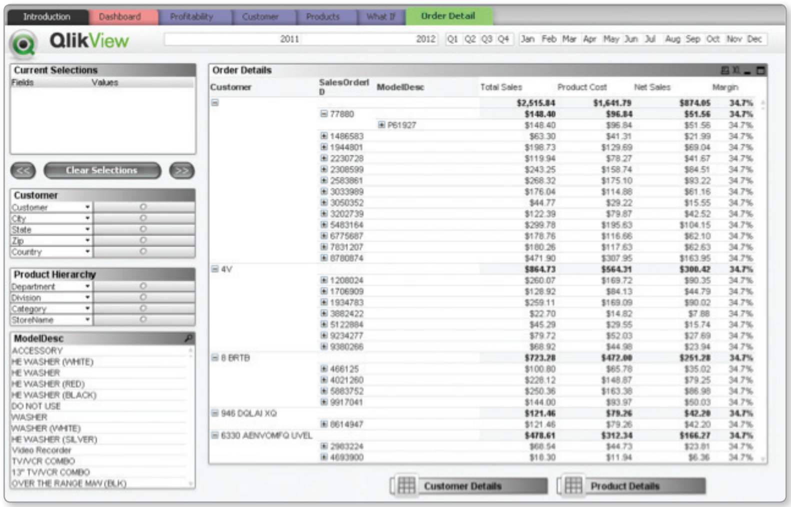
New Revenue : $ 2,242,749

Total Change In Revenue : $ 72,659

Department	Volume	Revenue	Margin	New Volume	New Revenue	New Margin	Revenue Change	Revenue Change %	Margin Change
Electronics	9,491	$360,563	$125,266	9,840	$372,635	$128,679	$12,072	3.3%	$3,41
Housewares and...	12,461	$467,212	$162,318	12,920	$482,855	$166,741	$15,643	3.3%	$4,42
Home Improve...	17,369	$657,024	$228,262	18,008	$679,022	$234,482	$21,999	3.3%	$6,22
Linens, Home D...	18,053	$685,291	$238,082	18,717	$708,236	$244,570	$22,945	3.3%	$6,48
	57,374	$2,170,090	$753,927	59,485	$2,242,749	$774,472	$72,659	3.3%	$20,54

- **Order Detail tab:** This tab provided order level analysis with total sales, product cost, net sales, and margin metrics. It also displayed the customer details such as address, city and state with product details down the SKU ID level. The selections available on the tab were; country, territory, store name, model description, country, city, state, zip code, department, division, category, year, quarter, and month fields (Figure 13).

Figure 13. Order Detail Tab of the QlikView Apps Used During Scalability Tests



ModelDesc

ACCESSORY
HE WASHER (WHITE)
HE WASHER
HE WASHER (RED)
HE WASHER (BLACK)
DO NOT USE
WASHER
WASHER (WHITE)
HE WASHER (SILVER)
Video Recorder
TV/VCR COMBO
13" TV/VCR COMBO
OVER THE RANGE MW (BLK)

Order Details

Customer	SalesOrderID	ModelDesc	Total Sales	Product Cost	Net Sales	Margin
⊟			**$2,515.84**	**$1,641.79**	**$874.05**	**34.7%**
	⊟ 77880		**$148.40**	**$96.84**	**$51.56**	**34.7%**
		⊞ P61927	$148.40	$96.84	$51.56	34.7%
	⊞ 1486583		$63.30	$41.31	$21.99	34.7%
	⊞ 1944801		$198.73	$129.69	$69.04	34.7%
	⊞ 2230728		$119.94	$78.27	$41.67	34.7%
	⊞ 2308599		$243.25	$158.74	$84.51	34.7%
	⊞ 2583861		$268.32	$175.10	$93.22	34.7%
	⊞ 3033989		$176.04	$114.88	$61.16	34.7%
	⊞ 3050352		$44.77	$29.22	$15.55	34.7%
	⊞ 3202739		$122.39	$79.87	$42.52	34.7%
	⊞ 5483164		$299.78	$195.63	$104.15	34.7%
	⊞ 6775687		$178.76	$116.66	$62.10	34.7%
	⊞ 7831207		$180.26	$117.63	$62.63	34.7%
	⊞ 8780874		$471.90	$307.95	$163.95	34.7%
⊟ 4V			**$864.73**	**$564.31**	**$300.42**	**34.7%**
	⊞ 1208024		$260.07	$169.72	$90.35	34.7%
	⊞ 1706909		$128.92	$84.13	$44.79	34.7%
	⊞ 1934783		$259.11	$169.09	$90.02	34.7%
	⊞ 3882422		$22.70	$14.82	$7.88	34.7%
	⊞ 5122884		$45.29	$29.55	$15.74	34.7%
	⊞ 9234277		$79.72	$52.03	$27.69	34.7%
	⊞ 9380266		$68.92	$44.98	$23.94	34.7%
⊟ 8 BRTB			**$723.28**	**$472.00**	**$251.28**	**34.7%**
	⊞ 466125		$100.80	$65.78	$35.02	34.7%
	⊞ 4021260		$228.12	$148.87	$79.25	34.7%
	⊞ 5883752		$250.36	$163.38	$86.98	34.7%
	⊞ 9917041		$144.00	$93.97	$50.03	34.7%
⊟ 946 DQLAI XQ			**$121.46**	**$79.26**	**$42.20**	**34.7%**
	⊞ 8614947		$121.46	$79.26	$42.20	34.7%
⊟ 6330 AENVOMFQ UVEL			**$478.61**	**$312.34**	**$166.27**	**34.7%**
	⊞ 2983224		$68.54	$44.73	$23.81	34.7%
	⊞ 4693900		$18.30	$11.94	$6.36	34.7%

SERVER CONFIGURATION

The server that was used for testing was an IBM x3650 server running Windows Server 2008 R2. The server had 192GB of RAM and was running sixteen 2.60 Ghz core CPUs. The server list price on the IBM web site was $16,289*. The system did not saturate with neither of the tests and QlikView Server was able to perform very well by only using some of the available hardware resources, including RAM and CPU. The hardware resources consumed during the tests are displayed below. The default QlikView Server configuration was used during the tests.

Figure 14. Hardware Resources Consumed During the Scalability Tests

Server Details	**Resources Used** **10 million app with 122,578 selections**	**Resources Used** **50 million app with 35,896 selections**
Available RAM: 192 GB	**RAM Used:** 136 GB	**RAM Used:** 149 GB
CPU: 162.60 Ghz cores	**Average CPU:** 39% Utilization	**Average CPU:** 52% Utilization

* The server list price on the IBM web site as of October 26, 2012

SESSION LENGTH

In real-world deployments, organizations have different types of users, from casual to power users. Depending on the user, their activity and session lengths vary. For these tests, we used different session lengths as the amount of time the user is connected to the system would vary based on their analysis types.

Users were split into 2 categories:

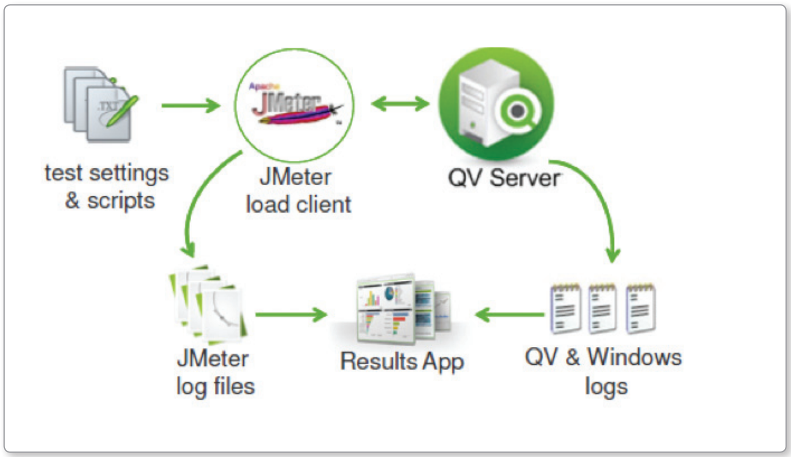
- Dashboard users: These users made less selections and had longer wait times between them
- Analyst users: These users were heavy users who made more selections and had shorter wait times

During the sessions, we simulated a variety of interactions, including applying selections, lassoing data in charts, opening tabs, manipulating sliders, applying what-if analysis, utilizing different aggregations on charts, basically any business discovery activities that a regular business user would perform.

Scalability Test Methodology

For this scalability testing, JMeter, which is a load/performance testing tool, was used to script the user interaction scenarios with different QlikView apps. For each scenario different usage patterns have been simulated towards the same environment (single medium server). The log files from the JMeter and QlikView Server were used to summarize the test results. Further details of the testing scenarios and the results are provided in the following section.

Figure 15. Scalability Testing Methodology



Scalability Results

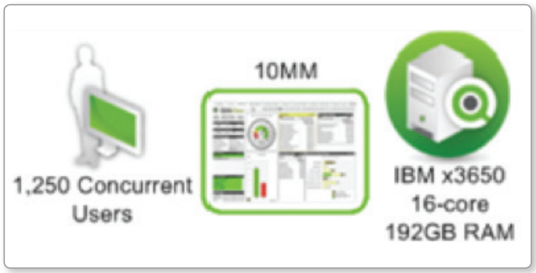
As we tested, we looked at a number of different metrics. These include:

- **Average Response Time:** This represents the average response time it took for QlikView Server to calculate the charts and display the information and the data associations for the selections.
- **Number of Selections:** This represents the total number of user interactions that are conducted during the test. User interactions include activities such as making selections on different data values, opening charts, manipulating sliders, applying what-if analysis, utilizing different aggregations on charts, and many more.
- **Total Sessions:** This is the number of total user sessions created during the test time frame.
- **Concurrent Users:** This represents the number of users who concurrently interacted with the QlikView app.
- **RAM Used:** This represents the memory usage by the QlikView Server during the test timeframe. Please note that most of this RAM was used by QlikView Server central cache. QlikView caches the calculations and data aggregations allowing QlikView chart calculations to be done once. Evidently the benefits are faster response times and lower CPU utilization.
- **Average CPU:** This represents the average CPU usage by the QlikView Server during the test timeframe.
- **Total Number of Rows:** This represents the total number of fact table rows that are analyzed on the QlikView apps.
- **User Types:** This represents the user types that are simulated during the test. The user types determined the session length and number of selections that were made by that user.
    - Dashboard User: These dashboard users made 2+ selections per minute and had active sessions of 7-8 minutes.
    - Analyst User: These analysis users made 3+ selections per minute and had active sessions of 10-12 minutes.

Small QlikView App Test

This test demonstrates the impact of a small QlikView app with 10 million rows. This is a real dashboard with 7 tabs, dozens of charts and detail tables for granular level data analysis.

Figure 16. Small QlikView App Test Setup

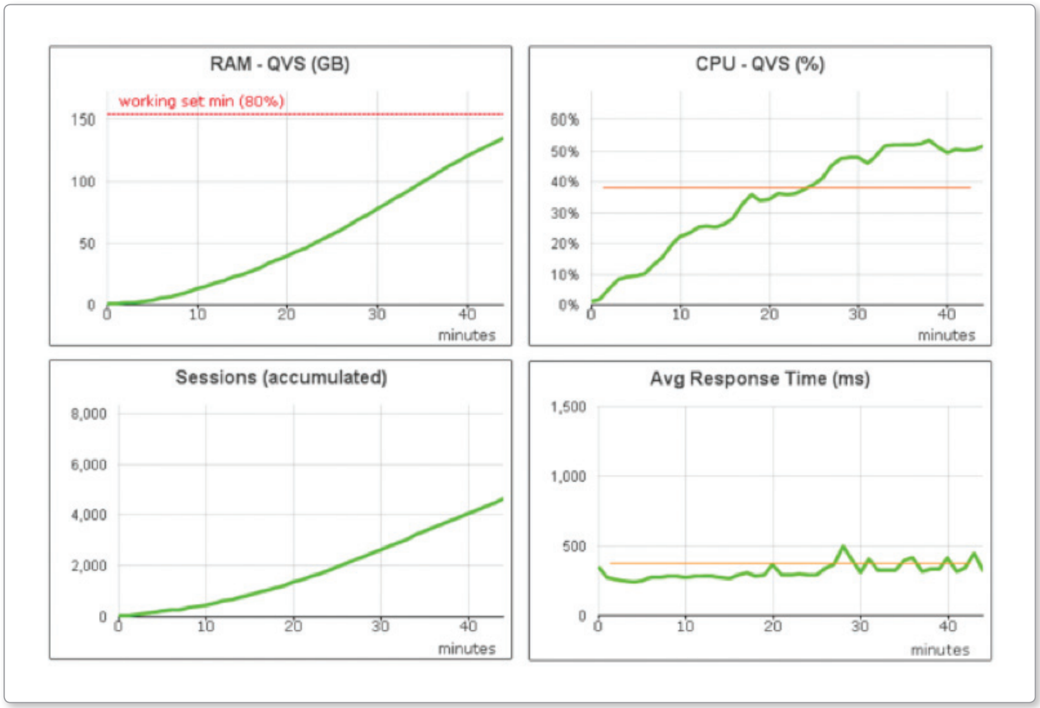


During the test, the simulation program created 4,647 sessions with 1,250 concurrent users. Constant user activity was maintained during the test was generating over 122,578 selections during 45 minutes. Total RAM usage on the server was 136 GB with average CPU usage of 39%. The average response time to calculate the charts and to display associations was 0.4 seconds.

Figure 17. Small QlikView App Performance Test Results

Test Scenario Details	Results
Number of server	1
Cost of server	$16,289*
Number of sessions	4,657
Number of concurrent sessions	1,250
Average response time	0.4 seconds
Number of selections generated	122,578

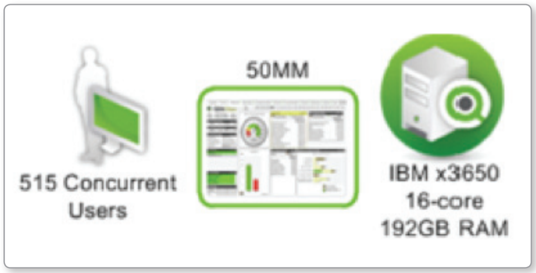
Figure 18. Small QlikView App Performance Test Results



Medium QlikView App Test

This test demonstrates the impact of a medium app of just over 50 million rows. This is a real dashboard with 7 tabs, dozens of charts and detail tables for granular level data analysis.

Figure 19. Medium QlikView App Test Setup

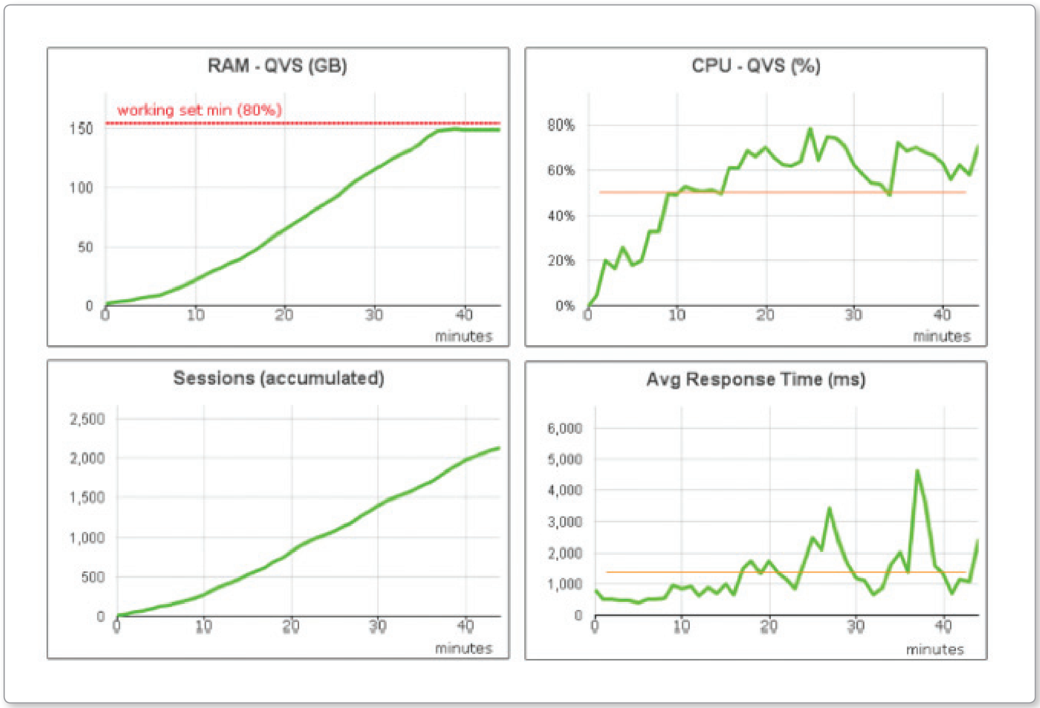


During the test, the simulation program created 2,139 user sessions with 515 concurrent users. Constant user activity was maintained during the test was generating over 35,896 selections during 45 minutes. Total RAM usage on the server was 149 GB with average CPU usage of 52%. The average response time to calculate the charts and to display associations was 1.4 seconds.

Figure 20. Medium QlikView App Performance Test Results

Test Scenario Details	Results
Number of server	1
Cost of server	$16,289*
Number of sessions	2,139
Number of concurrent sessions	515
Average response time	1.4 seconds
Number of selections generated	35,896

Figure 21. Medium QlikView App Performance Test Results



Conclusion

The scalability tests represented on this paper demonstrate that the scalability of a business intelligence tool should not be only measured against the data size or number of users by using a couple of reports. In real world, business users would be using tens of reports to get the insights they need. The scalability tests should measure real world analysis scenarios where the users would be asking streams of questions. That is why we have used QlikView apps with multiple tabs during these tests scenarios with dozens of charts. We wanted to simulate and demonstrate the BI throughput that can be achieved with QlikView on a single medium size server that only costs $16,289*.

During the tests, the program simulated business users lassoing data in charts and graphs, clicking on field values in list boxes, manipulating sliders, conducting what-if and market basket analysis and many more business discovery activities. With every user activity, QlikView Server was instantly updating all the data displayed in the apps around these selections by aggregating the data with different dimension combinations and by displaying the associative data representation.

With the testing summarized on this paper, we demonstrated how 1,250 **concurrent** users were able to ask and answer their own streams of questions. In 45 minutes, they generated 122,578 selections which would be equivalent to thousands of business questions only by using a single server. This is what we mean by business intelligence throughput. When we measured QlikView scalability, we did not limit our testing to a couple of reports, but we simulated thousands of users asking thousands of questions and getting answers in less than 0.4 seconds on average. Imagine what a cluster of QlikView Servers or a large QlikView Server could do with its linear scalability.

Reference Documents

QlikView Server Memory Management and CPU Utilization Technical Brief
http://www.qlikview.com/us/explore/resources/technical-briefs?language=english

QlikView Server Linear Scaling Technical Brief
http://www.qlikview.com/us/explore/resources/technical-briefs?language=english

Scaling Up vs. Scaling Out in a QlikView Environment Technical Brief
http://www.qlikview.com/us/explore/resources/technical-briefs?language=english

QlikView Architecture and System Resource Usage Technical Brief
http://www.qlikview.com/us/explore/resources/technical-briefs?language=english

QlikView Scalability Overview Technology White Paper
http://www.qlikview.com/us/explore/resources/whitepapers/qlikview-scalability-overview