

ReadSoft REPORTER

REPORTER 5-4 Performance FAQ

September 2011 - ReadSoft AB



© ReadSoft AB (publ). Licensees may make a number of copies, in paper form or in electronic form, of this material, corresponding to the number of allowed concurrent users of the software. The material may only be used in conjunction with operation of the software, by designated users, and in accordance with the Software License Agreement accompanying the software. Any other use, including but not limited to modification, translation and reproduction, and distribution is strictly forbidden.

The contents of this document are subject to change without notice.

ReadSoft is a registered trademark of ReadSoft AB. Other product and company names herein may be the trademarks or registered trademarks of their respective owners.

Questions or comments about this document may be emailed to documentation@readsoft.com.

12092011.0407

Contents

Introduction	4
REPORTER Performance FAQ	5
Quick References for Hardware and Performance1	6



Introduction

This document provides answers to the frequently asked questions related to the REPORTER solution.

Its purpose is to help ReadSoft personnel, primarily from pre-sales, consulting, and support organizations to successfully run REPORTER projects, correctly estimate the required hardware, and set the customer's expectations.

All answers are applicable to **REPORTER 5-4 Service Pack 1**, in which a number of significant performance improvements were implemented and where all required benchmark tests were performed. The correctness of the answers provided is not guaranteed for any future REPORTER releases, since new implemented functionality, bug fixes, and any new integrated QlikView releases may impact the performance of the solution and will require updates to this document.

Note that the document assumes prior knowledge of the architecture and basic terminology used in REPORTER.

Disclaimer

The benchmarking tests mentioned in this document were performed with a specific hardware setup and set of data. Although benchmark testing provides guidelines, results may vary due to differences in hardware and data content. As a consequence, the numbers listed in this document may vary in other systems.



REPORTER Performance FAQ

Question	Answer	
What can you say about the performance of REPORTER solution in general?	In general, the term 'performance' is very much overused. Typically, when someone asks about performance, they are normally asking about one of these five different (though related) areas:	
	1. Hardware requirements	
	a. What are the hardware requirements for REPORTER?	
	b. How much data volume can REPORTER effectively handle?	
	2. End-user performance when browsing reports	
	a. What is the typical response time that we can expect from the reports? How fast is the data updated once we make selections?	
	3. Full data reload performance	
	a. How long does it take to do a full data reload? Can we schedule it during the night to see fresh reports the next morning?	
	4. Data collection performance.	
	a. How does the REPORTER plug-in for INVOICES affects the production modules? Will my Interpret processing be slower?	
	5. Operational requirements	
	a. Are there any maintenance routines or other operational considerations that we need to be aware of to keep performance high?	
	These are just examples of typical questions. However, their	

These are just examples of typical questions. However, their variety shows the complexity of the question and the difficulty in providing a generic answer. Instead, it makes sense to focus on each area separately.



Question	Answer
What is the difference between REPORTER 5-4 Service Pack 1 and the base release of REPORTER 5-4 and how does it affect the performance of the solution?	 REPORTER 5-4 Service Pack 1 includes a new version of QlikView software (version 10.0) that significantly improves the overall performance of the system all by itself. In addition, multiple performance-related enhancements were made by our development team in the REPORTER application. Those changes led to the following performance-related improvements compared to the base release of REPORTER 5-4: a) REPORTER now reloads the full set of data 5 times faster! b) REPORTER now requires only half of the RAM that was previously used to process the same volume of data. c) REPORTER now uses all available CPU cores to perform calculations. d) REPORTER now requires only half of the space on the hard disk previously required for the QlikView data files. With these improvements in mind, this Service Pack is recommended for any installation that has a high volume of invoices or where there are issues with performance.
What is the maximum number of invoices that REPORTER can effectively handle?	Recent performance benchmark testing on REPORTER 5-4 Service Pack 1 confirms that it can handle systems that process up to five million invoices. Systems of this size require the appropriate hardware to run efficiently, however. Note that REPORTER is not limited to the mentioned five million invoices. It's just that the R&D team did not perform benchmark testing for volumes higher than that.
What are the main criteria when selecting hardware for the REPORTER application server?	Sufficient RAM and CPU power are the most essential parameters. Since REPORTER performs all calculations in the memory on the server, the demands put on RAM memory and CPU use increase with the amount of historical data to process.



Question	Answer	
How much RAM should we have on the server?	The amount of RAM that is required on the REPORTER server can be calculated based on the following considerations:	
	a) Aggregator.qvw requires 600MB for every 100.000 invoices (used by the nightly scheduled data reload task and qvb.exe processes).	
	b) ProcessAnalyser.qvw requires the same 600MB for every 100.000 invoices (used when users access reports with qvs.exe processes).	
	c) Extra 10% needs to be considered for every concurrent user logged in to REPORTER (used by qvs.exe processes).	
	Example: your annual volume is 500,000 invoices and you plan to be able to report for two years of data with a maximum of five concurrent users accessing reports simultaneously.	
	In this case, the total volume of invoices is 1,000,000 (500,000 * 2 years), so the minimum amount of RAM you will need on the server would be:	
	6 GB + 6 GB + (6GB * 10% * 5 concurrent users) = 15 GB	
	Rounding this up to 16 GB on the application server should be sufficient for up to five concurrent users to run REPORTER with one million invoices.	
	You can use the following matrix to estimate the amount of RAM required:	
	Volume (# of invoices in the database)5 concurrent users10 concurrent users	
	100.000 2 GB 4 GB	
	200.000 4 GB 6 GB	
	500.000 8 GB 12 GB	
	1.000.000 16 GB 24 GB	
	2.000.000 32 GB 48 GB	
	3.000.000 48 GB 64 GB	
	5.000.000 96 GB 128 GB	

Question	Answer	
What happens if I do not have enough RAM available in my application server for the volume of invoices I have?	When this happens, REPORTER will use virtual memory and start swapping to the hard disk. As a result, the performance of the REPORTER solution will degrade depending on and how much RAM you are missing and the speed of your hard disks.	
How many CPU cores should the server have?	 QlikView 10 is integrated into REPORTER 5-4 Service Pack 1 and utilizes all available CPU cores to perform calculations. Consequently, the more CPUs that are available, the better the performance will be. Our benchmark testing showed that upgrading from 2 to 4 CPUs contributed significantly to improve the performance achieved: the time it took to reload the full data reload was cut in half. You can use the following matrix to estimate the number of CPUs required: 	
	Volume (# of invoices in the database)5 concurrent users10 concurrent users100.0002 CPU cores2 CPU cores200.0002 CPU cores2 CPU cores200.0004 CPU cores4 CPU cores500.0004 CPU cores4 CPU cores1.000.0004 CPU cores8 CPU cores2.000.0008 CPU cores8 CPU cores3.000.00016 CPU cores16 CPU cores	



Question	Answer
Why is REPORTER so demanding on resources? I only have 1 million invoices in my system, and even though it is a large number, it does not seem to me to be so large that it should require that kind of hardware.	REPORTER collects a lot of various data associated with invoices. This is not only the invoice header data, but also includes line items, accounting codes, the entire audit trail, results of the field interpretation and verification, fields that are transferred from INVOICES, information about the occurred errors, etc. All of this ends up in 50-100 records in the database for each invoice, the total amount of information being 10-20 KB on average for a single invoice. To sum it up for your system with 1 million invoices, REPORTER effectively handles 50-100 million data rows with a total database size of 10-20 GB! And it is still capable of providing instant access to the data for several concurrent users without problems. Of course, volumes like this do require proper hardware to process them.
Should REPORTER be installed on a stand-alone server, or can it work side by side with other applications?	Technically, REPORTER can work side by side with the other applications (with the exceptions of some web-based software). However, a dedicated server is strongly recommended due to REPORTER's great demands on memory and the CPU.
Should we use a 64-bit or 32-bit server for REPORTER?	For best performance, always use the 64-bit server and 64-bit operating system for REPORTER. 32-bit servers should only be used for demo purposes. These systems have a built-in limit of 2GB of RAM, which is the maximum amount of memory that can be used by an application on a 32-bit platform. Since QlikView uses a considerable amount of computer memory, it will stop working once it reaches this limit (which will happen once the number of invoices in your system exceeds 100.000).



Question	Answer
Is there any specific recommendation for the size of the hard disk?	 The amount of the free space required on the hard disk can be estimated taking the following into consideration: a) Size of the software itself: 30 MB b) Size of temporary files (logs, etc.): allocate at least 20 GB c) Size of the QlikView data files: 3 GB to keep data for 1 million invoices Allowing 30 GB of free hard drive space on the REPORTER server is sufficient Note that this estimate does not take into account the size of the SQL databases that reside in the SQL Server. The space for those needs to be considered separately. Note: REPORTER can be configured to store copies of processed invoice images on the disk. If this is the case, make sure you allocate sufficient space for this storage. Example: Say that you process 500,000 three-page invoices annually, with an average image size of 50 Kb per page, and you plan to keep 2 years of historical data in your system. At least 150 GB will be required just for storing the images (500,000 x 3 x 50 Kb x 2 = 150 GB).
Can the server be set up as a virtual server?	 REPORTER can run in a virtual environment. However, for the best performance, it is recommended that you run REPORTER in a non-virtual environment for solutions with the large numbers of invoices. Compared to a physical server environment, the results of our benchmark testing shows that the speed of the full data reload operations drops by up to 30% in the virtual environment. Also, considering the amount of RAM and CPU required by the REPORTER solution, it might not be practical to run this virtually, especially when the infrastructure is shared and one host computer is shared by multiple virtual guests.
How fast does the network port need to be?	Standard TCP/IP network with 100 Mbps is a minimum requirement.



Question	Answer
Does REPORTER require a dedicated SQL Server?	Generally, no. Neither the REPORTER application itself nor QlikView use SQL Server heavily. The load on the SQL Server is minimal both during normal use of the REPORTER Web GUI and when the data is being fully reloaded, and as a result, there are no performance benefits from moving the SQL Server to a dedicated server. On most occasions it is practical to use the existing infrastructure and install REPORTER's databases on an existing SQL Server (the one that hosts the INVOICES database, for example).
Can SQL Server be installed on the same server as REPORTER application?	 Yes, but it is not recommended. It is advisable to separate SQL Server from the REPORTER application server in these situations: a) If your system is configured to import data from ICS/INVOICEIT/PROCESSIT systems. In these cases, the import agent imports the invoice data into the SQL Server and that work is primarily done using SQL Server resources. b) If your system is configured to use temporary tables for the INVOICES data. In this case, the import agent transfers the invoice data from the temporary tables into the permanent ones and that also occurs in SQL Server. c) Since QlikView demands so many resources, it tries to consume all available resources on the computer where it operates. This forces SQL Server to compete for the hardware resources negatively impacting the overall performance of the system.



Question	Answer
How much space should be allocated for REPORTER's databases?	 REPORTER uses three databases: Message Store database (named RS_RReporter by default) Configuration database (RS_RApprove) Workflow database (RS_RWorkflow) The Message Store database contains all of the report data. You need about 1.2 GB per 100,000 invoices. The other two are used by REPORTER Server to support operations such as user management, licensing, maintenance routines, and system activity, and they also contain important configuration parameters. Their growth rate is minimal, and it is sufficient to set them up with an initial database size of 20 MB, with a growth rate of 2 MB
Are there any specific operational tasks that need to be run periodically on the SQL Server to help maintain good performance of the REPORTER?	 Several operational procedures that are <i>not</i> automatically set up by REPORTER's installation or configuration programs <i>must</i> be performed (scheduled) regularly: Backup and disaster recovery plan. Database maintenance. Re-indexing of the tables. Control of the size of the database and availability of hard disk space. Deletion of old data (older than the reporting scope) from the Message Store database. See the REPORTER Installation Guide for more information on these tasks ReadSoft does not take responsibility for problems that may occur due to failure to perform such tasks on a regular basis.



Question	Answer	
How long does it take to reload data and make it available in the reports? Is there enough time to get this done during night hours?	The reload job is executed by QlikView on the REPORTER application server and can be done during the night for databases with up to 5 million invoices (providing the appropriate hardware is used). Benchmark testing was performed on a medium performance server (Virtual server, 4 CPU, 12 GB RAM) and resulted in the following numbers:	
	Database size:	Full data reload time:
	100.000 invoices	00:10:36
	200.000 invoices	00:19:29
	300.000 invoices	00:29:20
	500,000 invoices	00:48:17
	900,000 invoices	01:25:02
	1,000,000 invoices	01:29:59
	Database size:1.000.000 invoices3.000.000 invoices5.000.000 invoices	Full data reload time: 01:05:30 03:16:08 06:40:00
How does the installation of the REPORTER Plug-in for INVOICES affect the performance of existing INVOICES installation?	This is a data collection plug-in a computer where one or more IN For small and medium installation no effect on existing INVOICES INVOICES operators do not even Large installations, on the other significantly, and specific invoice interpretation) may become 20-3 To remove the impact of this, yo the feature to use temporary table the REPORTER Configuration T to do this.	that is installed on each VOICES modules are installed. ons, the plug-in has practically installation, and in most cases, on notice it. hand, may be affected more e operations (such as 30% slower. ou should activate and configure es for the INVOICES data. See Fool Help file for details on how
Should we upgrade the hardware for computers where INVOICES modules are installed?	No. The REPORTER plug-in do to the RAM, CPU, or hardware n INVOICES.	es not add any significant load resources that are used by

Question	Answer	
How does the REPORTER plug- in for INVOICES affect the total load on the network?	Total load generated by the REPORTER plug-in for INVOICE modules is relatively small and will hardly be noticable in any network environment.	
Are there any performance implications when the REPORTER plug-in is run under Citrix?	No. Both INVOICES and the REPORTER plug-in for INVOICES run well under Citrix, with no limitations.	
What is the hardware recommendation for computers where users are running REPORTER in a web browser?	To optimize performance, REPORTER loads portions of the data that is selected by the user into the computer's memory. To avoid performance degradation, we recommend that the computers have a minimum of 1 GB of RAM. There are no other specific hardware requirements.	
What is the average response time that I can expect from the reports when I do basic operations, such as opening new reports, making selections, filtering and sorting charts and tables, for example?	 For most installations, the response time is instant and is within 1-10 seconds depending on the selection and the type of the report. For very high volumes of invoices, the response time is slightly slower. Recent benchmarking was conducted by R&D during a Production Simulation Test in which several concurrent users were on a system with 3 million invoices. 	
	 These were the response times during this test: Opening reports – within 15-20 seconds Making basic selections – 10-15 seconds Sorting and filtering tables and charts – 1-3 seconds Launching complex reports, such as performance or accruals reports that include pivot tables with massive amounts of data – 30-50 seconds 	
	Some of these numbers may seem higher than you would expect, however, note that they are for 3 million invoices and several concurrent users. In addition, the hardware used for this test was lower than what is recommended (only 16 GB of RAM was used on the server when 48 GB is recommended for five concurrent users).	
	So the response time for smaller systems or systems with hardware that matches the recommendations will be better than those numbers.	

Question	Answer
Is it possible to receive more information about the performance test benchmarks that were done at R&D?	Yes, the performance report with the summary of the tests done and their results is available on ReadSoft's internal SharePoint site in the downloads section for the REPORTER product.



Quick References for Hardware and Performance

The following table provides a quick reference to hardware requirements and performance figures.

Hardware specifications	Recommendation		
RAM on computer where REPORTER is installed	Use this table to estimate Volume (# of invoices in the database) 100.000 200.000 500.000 1.000.000 3.000.000 5.000.000 See page 7 for more infor	the amount of RAM r 5 concurrent users 2 GB 4 GB 8 GB 16 GB 32 GB 48 GB 96 GB	required: 10 concurrent users 4 GB 6 GB 12 GB 24 GB 48 GB 64 GB 128 GB
No. of CPUs on computer where REPORTER is installed	Use this tabe to estimate Volume (# of invoices in the database) 100.000 200.000 500.000 1.000.000 2.000.000 5.000.000 See page 8 for more information	the number of CPUs r 5 concurrent users 2 CPU cores 2 CPU cores 4 CPU cores 4 CPU cores 8 CPU cores 8 CPU cores 16 CPU cores 16 CPU cores	equired: 10 concurrent users 2 CPU cores 2 CPU cores 4 CPU cores 4 CPU cores 8 CPU cores 8 CPU cores 16 CPU cores
Server type	64-bit server (see page 9 for more information).		

Free hard disk space	30 GB + space for SQL databases + space for copies of invoice images if desired (see page 10 for more information).			
Network-port speed	Standard TCP/IP network with 100 Mbps minimum (see page 10).			
Dedicated SQL Server	No (see page 11).			
SQL on same server as REPORTER	No (see page 11).			
Space for databases	1.2 GB per 100,000 invoices (see page 12).			
Minimum RAM running REPORTER in a web browser	1 GB. No other hardware requirements. (see page 14).			
Performance issue	What can be expected			
Expected reload time for data	The following examples provide information on what can be expected.Using a virtual server with 4 CPUs and 12 GB RAM): $\boxed{\text{Database size:}}$ $\boxed{\text{Full data reload time:}}$ $100,000$ invoices $00:10:36$ $200,000$ invoices $00:19:29$ $300,000$ invoices $00:29:20$ $500,000$ invoices $00:48:17$ $900,000$ invoices $01:25:02$ $1,000,000$ invoices $01:29:59$ Using a physical server with 4 CPUs and 16 GB RAM: $\boxed{\text{Database size:}}$ $\boxed{\text{Full data reload time:}}$ $1.000,000$ invoices $01:05:30$ $3.000.000$ invoices $03:16:08$ $5.000.000$ invoices $06:40:00$ See page 13 for more information.			

