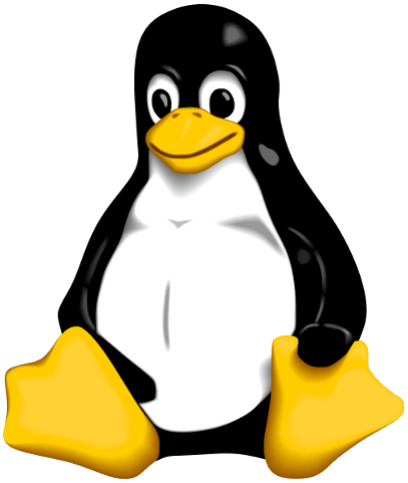


Design of Experiment

Igor Ljubuncic aka Dedoimedo

www.dedoimedo.com

Contents

List of Figures

About

Dedoimedo (www.dedoimedo.com) is a website specializing in step-by-step tutorials intended for human beings. Everything posted on my website is written in plain, down-to-Earth English, with plenty of screenshot examples and no steps ever skipped. You won't easily find tutorials simpler or friendlier than mine.

Dedoimedo lurks under the name of Igor Ljubuncic, a former physicist, currently living the dream and working as a Linux Systems Expert, hacking the living daylight out of the Linux kernel. Few people have the privilege to work in what is essentially their hobby and passion and truly love it, so I'm most grateful for the beauty, freedom and infinite possibilities of the open-source world. I also hold a bunch of certifications of all kinds, but you can read more about those on my website.

Have fun!

Copyright

This document is available under following conditions:

It is free for personal and education purposes. Business organizations, companies and commercial websites can also use the guide without additional charges, however they may not bundle it with their products or services. Said bodies cannot sell or lease the guide in return for money or other goods. Modifications are not permitted without an explicit approval from the author. All uses must be accompanied with credits and a link to www.dedoimedo.com.

You may also mirror and hotlink to this document. You must credit me for any such use.

In all eventualities, Dedoimedo retains all rights, explicit and implicit, to the original material. The copyright section may change at any time, without prior notice. For any questions, please contact me by email.

Disclaimer

I am not very fond of disclaimers, but they are a necessary part of our world.

So here we go:

I must emphasize the purpose of this guide is educational. It is not an official document and should not be treated as such. Furthermore, I cannot take any responsibility for errors, inaccuracies or damages resulting from the use of this book and its contents.

All of the material in this guide has been carefully worded and prepared. However, if for some reason you may feel it infringes on copyright or intellectual property of another work, please contact me with a detailed explanation pointing to the troublesome parts and I will try to sort the problem in the best way possible.

For any news or updates, you should always refer first to www.dedoimedo.com.

1 Introduction

Here are some very important points that must be taken into consideration before such an experiment is conducted – and more importantly – analyzed.

1.1 Setup

It is important to note that many other factors can affect the results of such an experiment, including the speed of the hard disks, the age of the computer, the quantity of real-time software, and many more. Nevertheless, we will perform the experiment, under the following restrictions:

- Windows XP SP2 operating system with all the updates installed.
- Operating system installed just before the experiment.
- No other real-time software installed except ProcessExplorer.
- Classic Windows theme, screen resolution 800x600.
- 4GB C: partition (NTFS) with 65% free space.
- 2GB secondary partition (NTFS) with 95% free space.
- Temporary folders, caches and the Recycle Bin are empty.
- System Restore is disabled.
- Hardware configuration.
- AMD Athlon 64 3800+ processor (single core used) .
- VMware IDE hard drive (on top of physical WD2500KS 250GB HDD).

1.2 Software configuration

The following software was used for the experiment:

- Grisoft AVG 7.5.503 anti-virus (free for personal use).
- McAfee VirusScan anti-virus (free with subscription, limited time offer) .
- Sygate Personal Firewall 5.6.2808 (free, discontinued, still available).
- Comodo Personal Firewall2.4 (free for personal use).

2 Benchmarking

Two responses will be measured:

Boot up time Will be measured manually by stopwatch, from cold start to loading of last processes, using ProcessExplorer; while there are more accurate benchmarking tools available, this method will suffice.

Performance with security software running in the background This will be determined by the scan time of another on-demand only security program Safer Networking Spybot Search & Destroy 1.5. Alternatively, the performance can also be determined by searching for a file, copying of files, defragmentation of a partition, etc.

3 Design of Experiment

3.0.1 Questions

What is the optimal combination of anti-virus and firewall for a 256MB Windows XP machine? What is the optimal combination for a 512MB machine? What is the deciding performance factor – the hardware or the software?

3.1 Objective

Define the best combination of RAM and security software for low-end PCs.

3.2 Responses

Boot up time and the overall responsiveness of the operating system are the two most critical parameters to most users. While other factors could be benchmarked as well, we shall limit our experiment to the above two.

3.3 Input factors

Most Windows users run their systems protected by a basic combination of a software firewall (quite often the built-in Windows firewall, not tested here) and an anti-virus

product (usually a preinstalled brand name). While such setups could be considered adequate from the security aspect, given the right behavior, they impact the user's experience in more than just the protection they provide, namely the performance of the operating system and installed applications is affected. In this experiment, we shall examine the impact of three factors – one hardware and two software – on the responsiveness of the system.

- Hardware: the size of RAM (low-spec 256MB, high-spec 512MB).
- Software: anti-virus (AVG, McAfee) and firewall (Sygate, Comodo).

Based on the web rumors, the two programs in each category represent a "light" and a "heavy" product, although the actual performance has yet to be tested.

3.4 Factor levels

As said, we shall base our levels on the existing reputation of the said products (save RAM, where the levels are self-evident).

- AVG – as the low-end (-1) anti-virus
- McAfee – as the high-end (+1) anti-virus
- Sygate – as the low-end (-1) firewall
- Comodo – as the high-end (+1) firewall

3.5 Experimental design

Since our experiment revolves around 3 factors with 2 levels, we shall perform a 2k full factorial design (a total of 8 runs).

3.6 Replicates

Single series only.

3.7 Experiment results

Table 1: Design of Experiment results

Runs	RAM size (MB)	Anti-virus	Firewall	Boot time (sec)	Scan time (sec)
1	256	AVG	Sygate	89	861
2	256	AVG	Comodo	98	921
3	256	McAfee	Sygate	108	954
4	256	McAfee	Comodo	121	964
5	512	AVG	Sygate	86	857
6	512	AVG	Comodo	80	812
7	512	McAfee	Sygate	107	863
8	512	McAfee	Comodo	106	929

4 Analysis

We have analyzed the results using Minitab 14.

4.1 Boot up time

In the first run, we have included all the interactions. We can see that this model is too complex and must be reduced.

Figure 1: Boot time versus RAM, Anti-virus and Firewall

Factorial Fit: Boot up time versus RAM, Anti-virus, Firewall

Estimated Effects and Coefficients for Boot up time (coded units)

Term	Effect	Coef
Constant		99.375
RAM	-9.250	-4.625
Anti-virus	22.250	11.125
Firewall	3.750	1.875
RAM*Anti-virus	1.250	0.625
RAM*Firewall	-7.250	-3.625
Anti-virus*Firewall	2.250	1.125
RAM*Anti-virus*Firewall	0.250	0.125

S = *

Analysis of Variance for Boot up time (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	1189.38	1189.38	396.458	*	*
2-Way Interactions	3	118.38	118.38	39.458	*	*
3-Way Interactions	1	0.13	0.13	0.125	*	*
Residual Error	0	*	*	*		
Total	7	1307.88				

We can see that the 3-way interaction is the least significant and shall remove it from the analysis.

Figure 2: Three-way interaction removed

Factorial Fit: Boot up time versus RAM, Anti-virus, Firewall

Estimated Effects and Coefficients for Boot up time (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		99.375	0.1250	795.00	0.001
RAM	-9.250	-4.625	0.1250	-37.00	0.017
Anti-virus	22.250	11.125	0.1250	89.00	0.007
Firewall	3.750	1.875	0.1250	15.00	0.042
RAM*Anti-virus	1.250	0.625	0.1250	5.00	0.126
RAM*Firewall	-7.250	-3.625	0.1250	-29.00	0.022
Anti-virus*Firewall	2.250	1.125	0.1250	9.00	0.070

S = 0.353553 R-Sq = 99.99% R-Sq(adj) = 99.93%

Analysis of Variance for Boot up time (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	1189.38	1189.38	396.458	3171.67	0.013
2-Way Interactions	3	118.38	118.38	39.458	315.67	0.041
Residual Error	1	0.13	0.13	0.125		
Total	7	1307.88				

The RAM-Anti-virus interaction is not significant ($P > 0.05$); we shall remove it.

Figure 3: Boot time anti-virus interaction is insignificant

Factorial Fit: Boot up time versus RAM, Anti-virus, Firewall

Estimated Effects and Coefficients for Boot up time (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		99.375	0.4507	220.49	0.000
RAM	-9.250	-4.625	0.4507	-10.26	0.009
Anti-virus	22.250	11.125	0.4507	24.68	0.002
Firewall	3.750	1.875	0.4507	4.16	0.053
RAM*Firewall	-7.250	-3.625	0.4507	-8.04	0.015
Anti-virus*Firewall	2.250	1.125	0.4507	2.50	0.130

S = 1.27475 R-Sq = 99.75% R-Sq(adj) = 99.13%

Analysis of Variance for Boot up time (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	1189.38	1189.38	396.458	243.97	0.004
2-Way Interactions	2	115.25	115.25	57.625	35.46	0.027
Residual Error	2	3.25	3.25	1.625		
Total	7	1307.88				

Again, we shall reduce our model further, by removing the Anti-virus-Firewall interaction.

Figure 4: Anti-virus firewall interaction removed

Factorial Fit: Boot up time versus RAM, Anti-virus, Firewall

Estimated Effects and Coefficients for Boot up time (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		99.375	0.7465	133.12	0.000
RAM	-9.250	-4.625	0.7465	-6.20	0.008
Anti-virus	22.250	11.125	0.7465	14.90	0.001
Firewall	3.750	1.875	0.7465	2.51	0.087
RAM*Firewall	-7.250	-3.625	0.7465	-4.86	0.017

S = 2.11148 R-Sq = 98.98% R-Sq(adj) = 97.61%

Analysis of Variance for Boot up time (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	1189.38	1189.38	396.458	88.93	0.002
2-Way Interactions	1	105.13	105.13	105.125	23.58	0.017
Residual Error	3	13.37	13.37	4.458		
Total	7	1307.88				

This is our final, reduced model. Although P $>$ 0.05 for the Firewall factor, we cannot remove it from the equation, as the interaction between the RAM and the Firewall is significant. You can see the results below, in the graphical form.

Figure 5: Residual plots for boot time

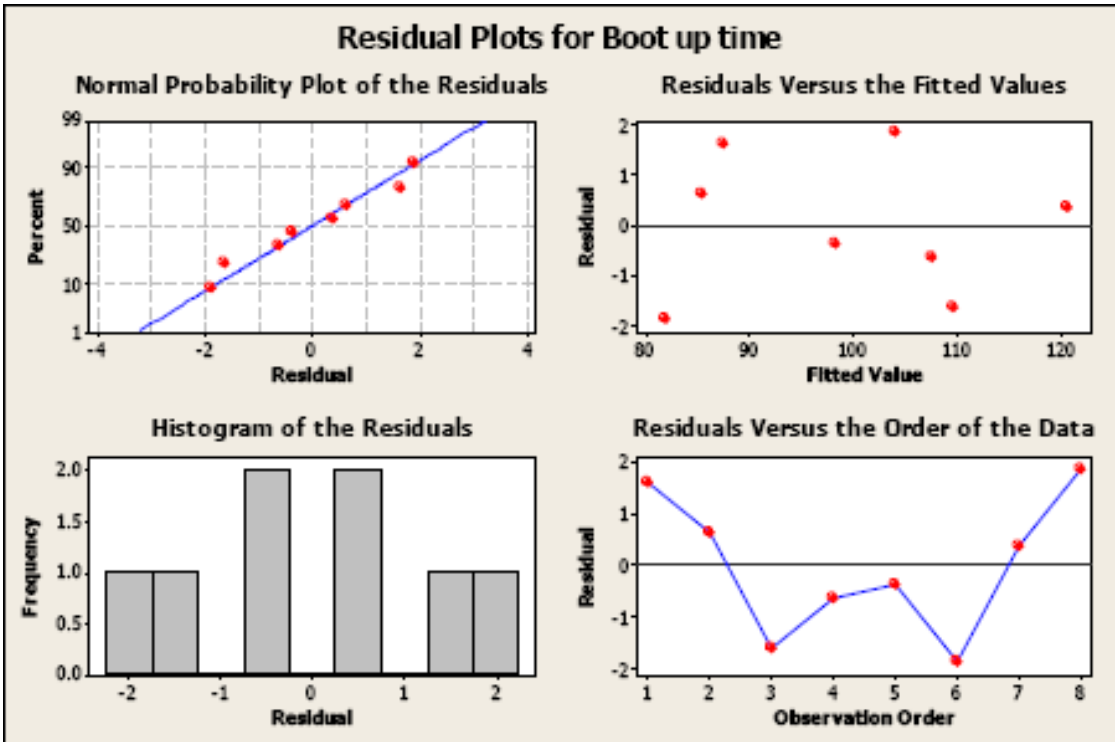


Figure 6: Boot time normal probability plot of the standardized effects

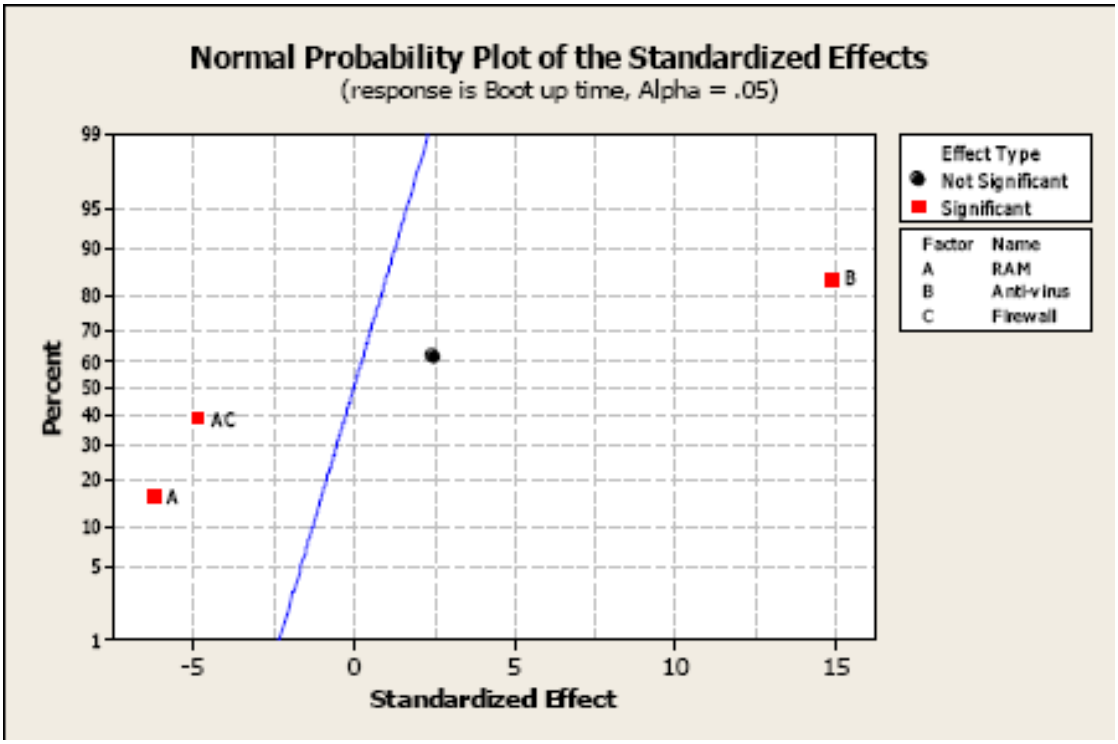
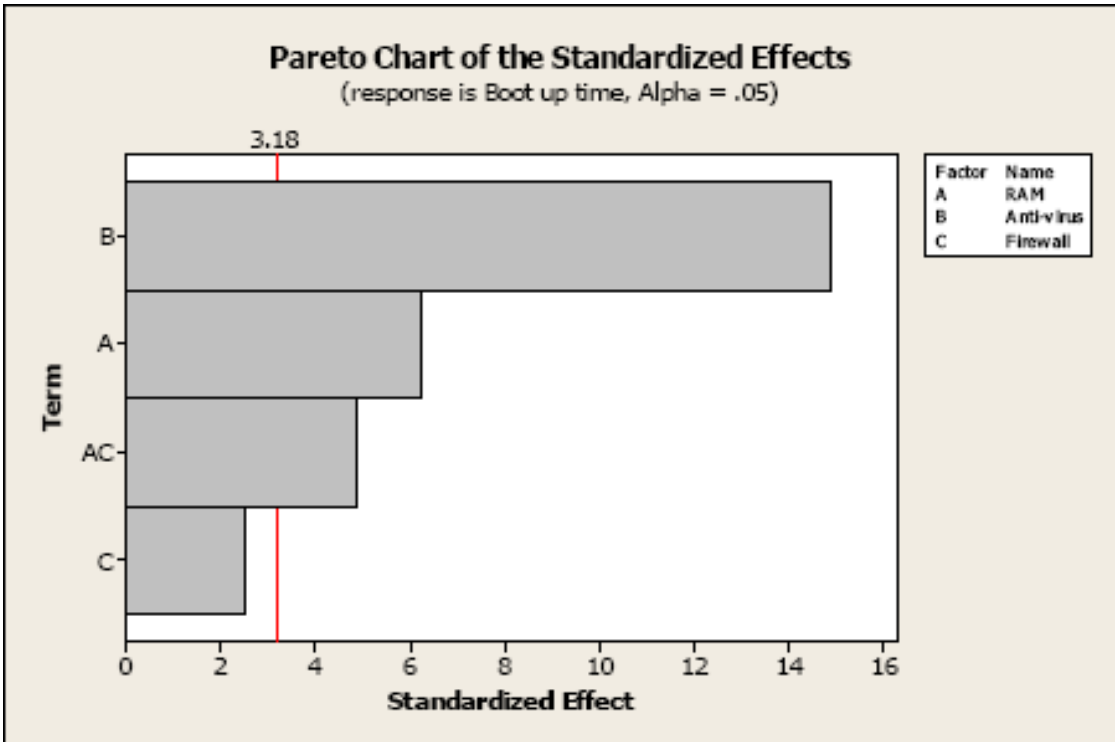
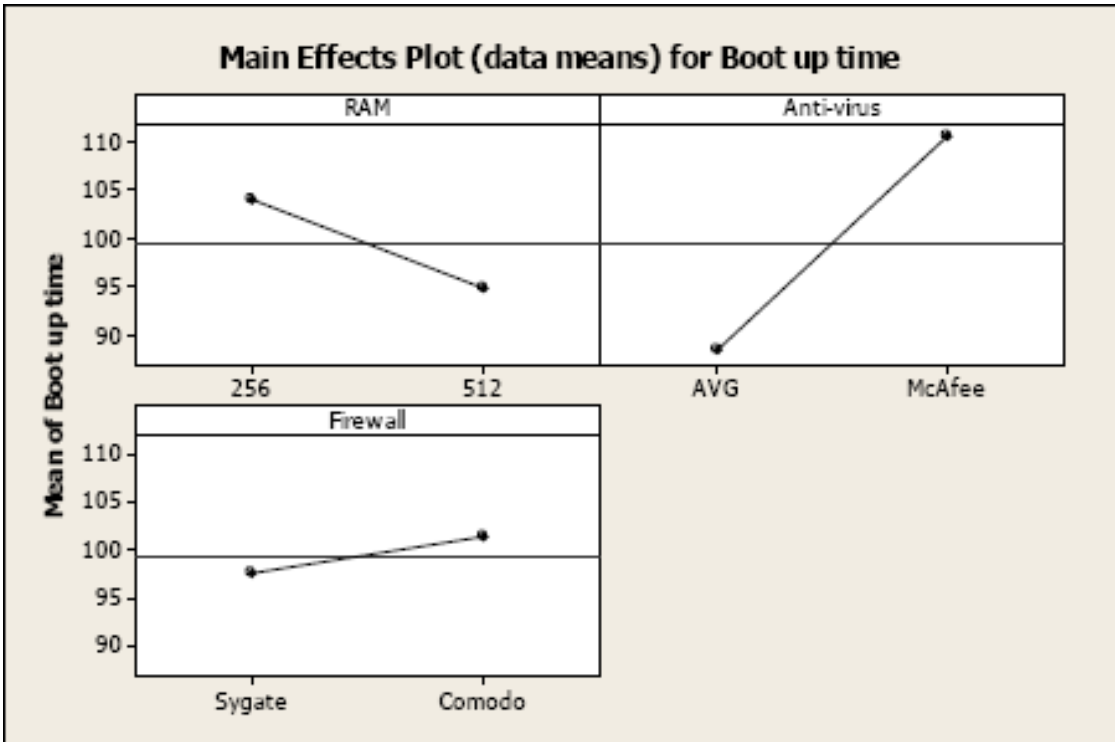


Figure 7: Boot time Pareto chart of the standardized effects



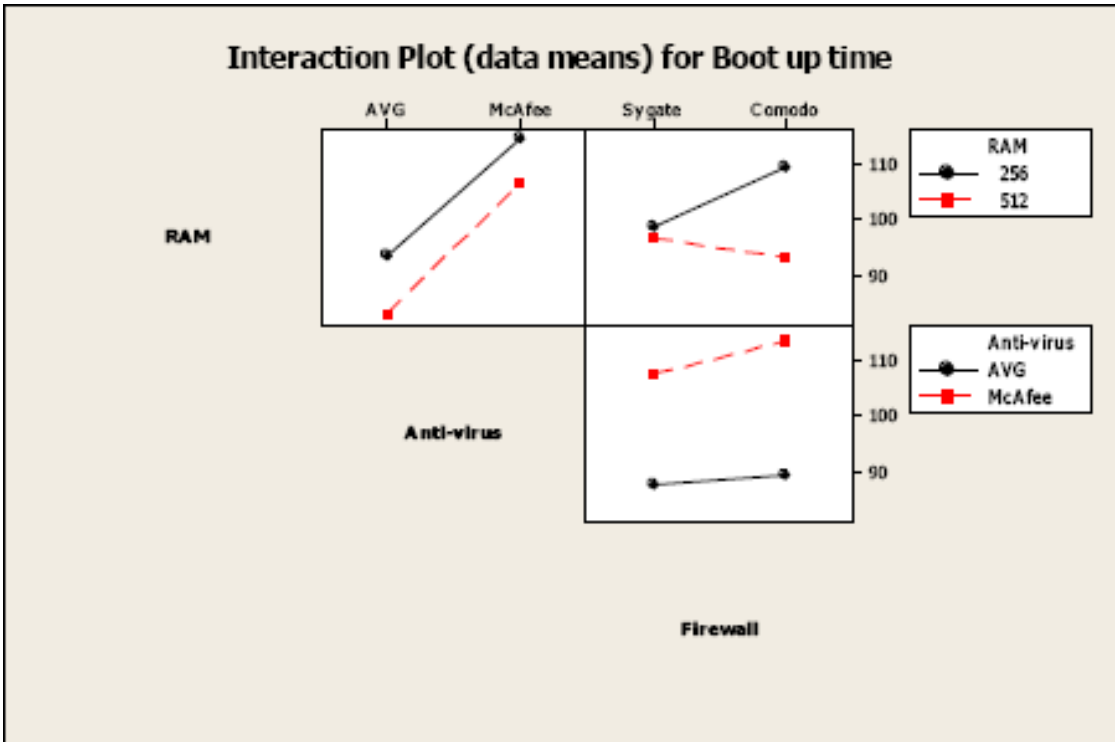
The most significant factor is the anti-virus – rather than RAM – which is a somewhat surprising (yet encouraging) fact for PC users with low-end machines. Furthermore, our experiment is correct in 97.61% (R2 adj) of cases. Now, let's examine the main effects for the Boot up time:

Figure 8: Main effects plot for boot time



The most significant factor is indeed the Anti-virus, followed by the RAM. In both cases, the experiment results agree with the popular definition of a "light" versus a "heavy" product. For the firewall, there was little difference overall, with a slight advantage in favor of the "light" product.

Figure 9: Interaction plot for boot time



We can see that the boot up process is linear for both the RAM size and the anti-virus product, with some interaction for the firewall. This might be explained by the following facts: The last version of the Sygate firewall was produced in 2004 and has not been updated since. In 2004, most computers ran on very little RAM, mainly 256MB. On the other hand, the Comodo firewall is a new and constantly updated program, optimized for new machines with powerful processors and a plenty of RAM. We shall examine the differences between the firewalls at a greater depth later.

4.1.1 Recommendation to users

The choice of the anti-virus program is the most important factor – of the three examined in this experiment – determining the boot up time of a PC running Windows XP operating system, with the memory size only in the second place, contrary to the intuition. This means that even the users of low-end machines can achieve reasonable performance with the right choice of an anti-virus product. The choice of a firewall is less important, although Sygate favors machines with less RAM. Indeed, Sygate has been known as one of the lightest product available on the market.

4.2 Scan time

Again, using the same principles, we have analyzed the second parameter, the scan time of the drive C: using Spybot Search & Destroy anti-malware program. Below is the reduced model, without the 2-way interactions.

Figure 10: Scan time two-way interaction results

Factorial Fit: Scan time versus RAM, Anti-virus, Firewall

Estimated Effects and Coefficients for Scan time (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		895.13	5.724	156.38	0.000
RAM	-59.75	-29.87	5.724	-5.22	0.014
Anti-virus	64.75	32.38	5.724	5.66	0.011
Firewall	22.75	11.38	5.724	1.99	0.141
RAM*Anti-virus*Firewall	40.25	20.13	5.724	3.52	0.039

S = 16.1903 R-Sq = 96.18% R-Sq(adj) = 91.09%

Analysis of Variance for Scan time (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	16560.4	16560.4	5520.1	21.06	0.016
3-Way Interactions	1	3240.1	3240.1	3240.1	12.36	0.039
Residual Error	3	786.4	786.4	262.1		
Total	7	20586.9				

You may notice that the 2-way interactions are insignificant, however the 3-way interaction remains strong with a low P value – whereas the firewall factor is not a significant contributor. This presented us with a logical problem. Therefore, we have decided to reduce the problem to only the RAM and Anti-virus factors, as they are the two major contributors.

Figure 11: Scan time versus RAM and Anti-virus

Factorial Fit: Scan time versus RAM, Anti-virus

Estimated Effects and Coefficients for Scan time (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		895.13	11.25	79.57	0.000
RAM	-59.75	-29.87	11.25	-2.66	0.045
Anti-virus	64.75	32.38	11.25	2.88	0.035

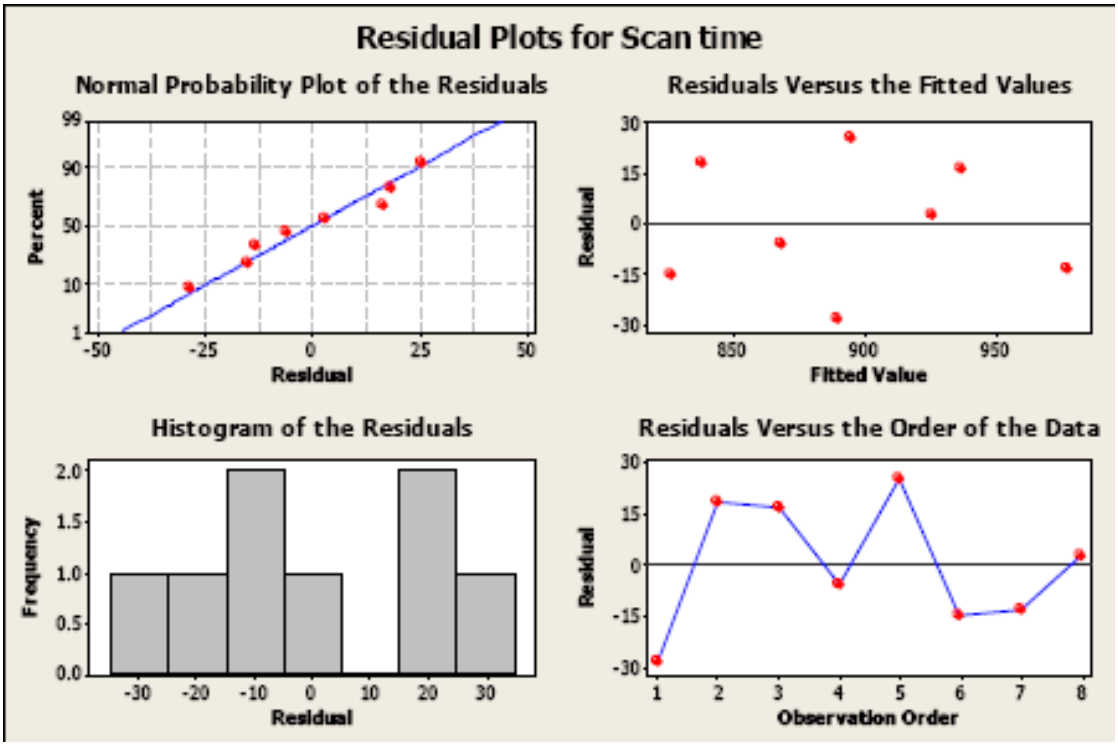
S = 31.8171 R-Sq = 75.41% R-Sq(adj) = 65.58%

Analysis of Variance for Scan time (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	2	15525.3	15525.3	7762.63	7.67	0.030
Residual Error	5	5061.6	5061.6	1012.33		
Lack of Fit	1	21.1	21.1	21.12	0.02	0.903
Pure Error	4	5040.5	5040.5	1260.13		
Total	7	20586.9				

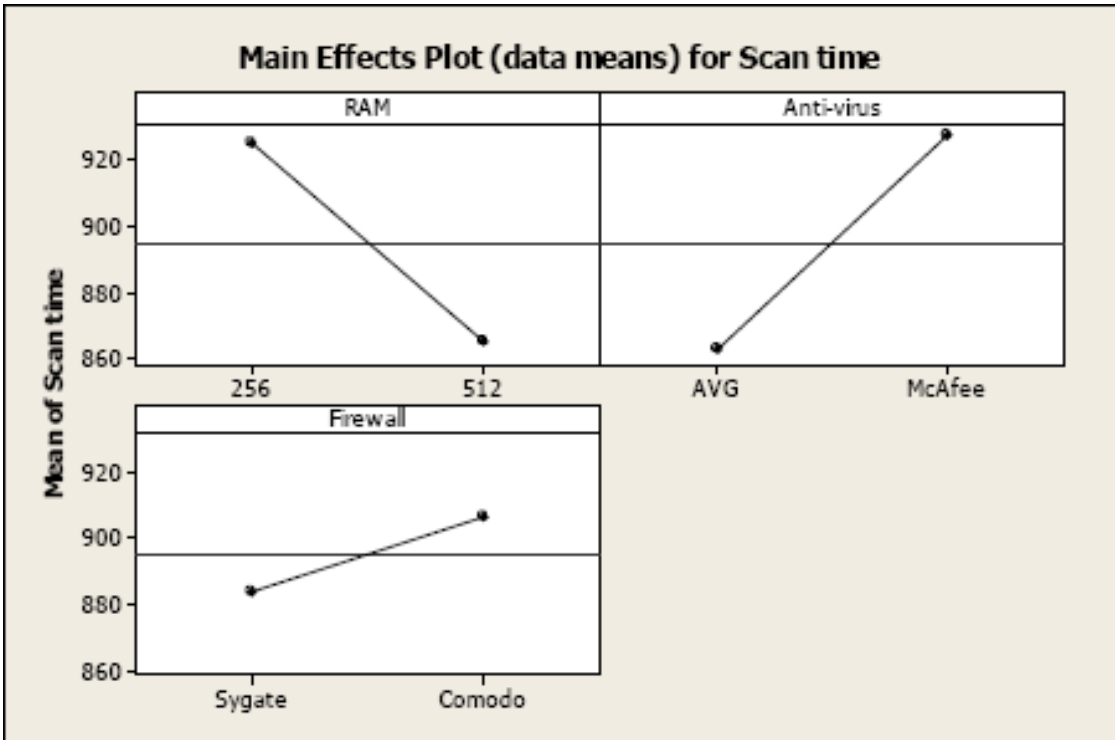
However, this reduced model can explain only 65.58% of the cases, compared to 91.09% for the previous model. Risking some possible misunderstanding of the model, which we shall try to explain separately, we will proceed with the original reduced mode, containing the 3-way interaction without the 2-way interactions. Again, similarly to the boot up time results, we can see that the anti-virus is the most significant factor, followed by the RAM size. The choice of a firewall did not affect the results.

Figure 12: Scan time residual plots



We can see that the distribution of residuals is not normal, compared to the boot up time, indicating that the scan time is most likely a nonlinear process, which can possibly explain the convoluted results. An experiment with center points might be in order to verify this.

Figure 13: Scan time main effects plot



Using the AVG anti-virus instead of the McAfee one will reduce the scan time drastically. The same applies to the RAM size.

Figure 14: Scan time interaction plot

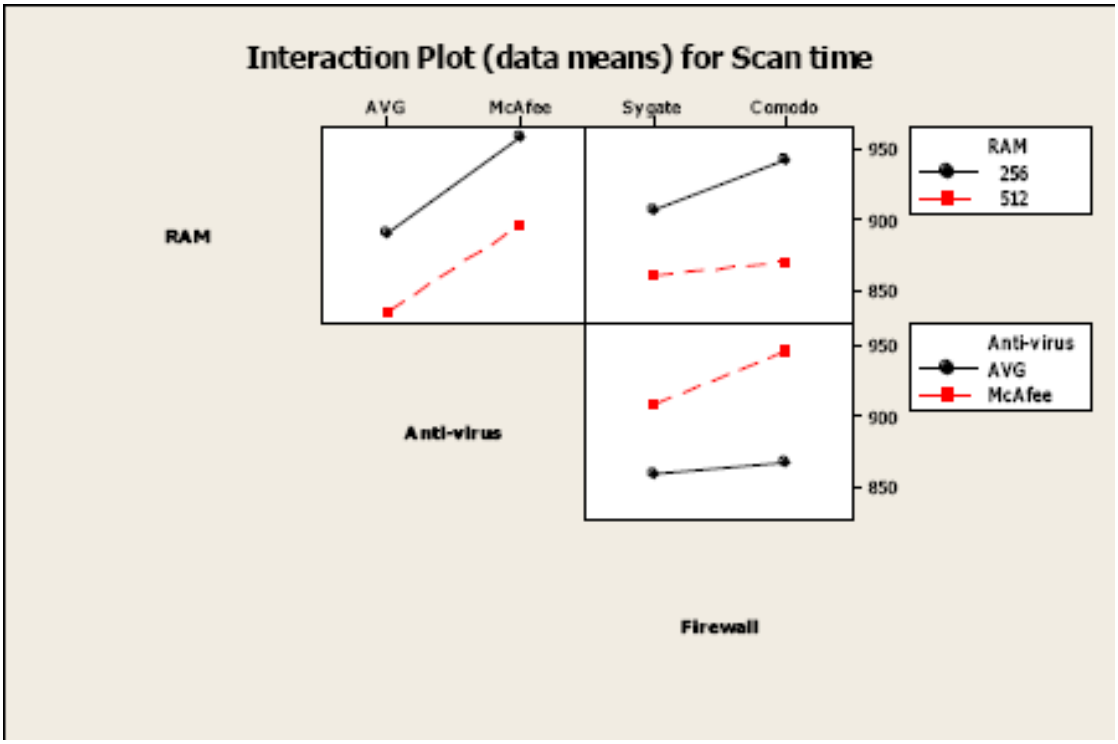
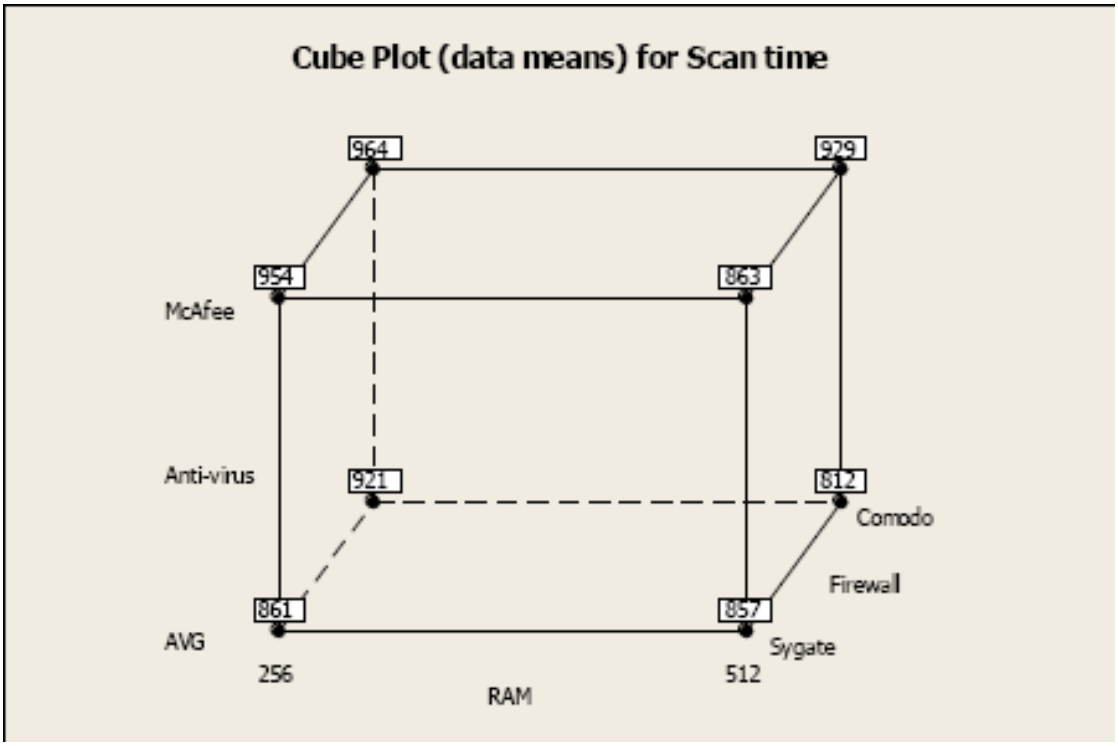


Figure 15: Scan time cube plot



The combination of the RAM and the anti-virus did not change much. This means that the minimum scan time is probably limited by an unknown factor that was not measured in the experiment, which might be the processor clock or the speed of the hard disk. In other words, the optimal scan time can be achieved by either the light choice of the anti-virus or the RAM size, but not necessarily both. This is encouraging for the Windows users with only 256MB RAM. The firewall choice only slightly affected the results, in the favor of Sygate, which is the lighter product, designed in the age of slower machines with less RAM. Furthermore, Comodo firewall does install additional drivers monitoring extra processes, in addition to the TCP/IP stack, which could partially explain the 3-way interaction, while not directly contributing to any of the 2-way interactions.

4.2.1 Recommendation to users

The choice of the anti-virus program is the most important factor determining the scan time, with the memory size in the second place. This means that even the users of low-end machines can achieve performance comparable to more powerful computers by choosing a "light" anti-virus product.

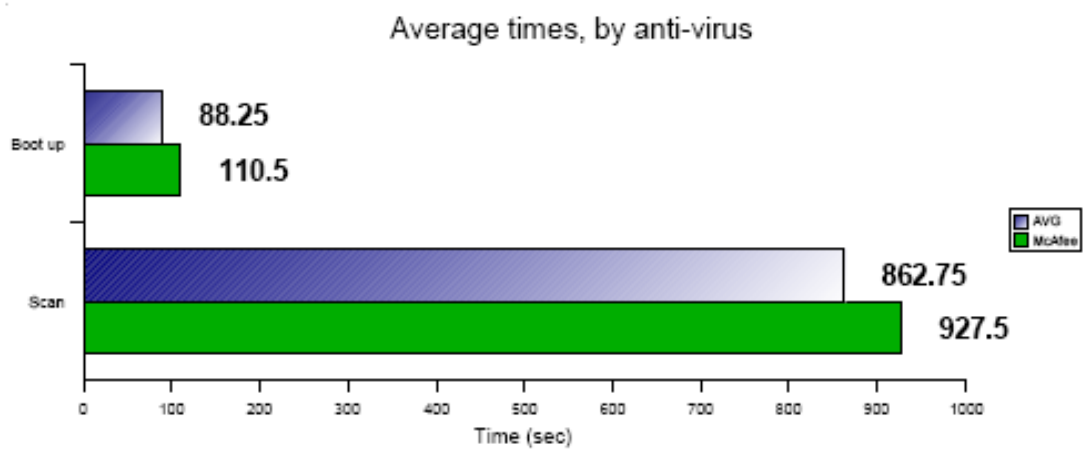
4.3 Average performance differences

Let's see how much the average PC user will gain by choosing this or that product.

4.3.1 Anti-virus

Using McAfee anti-virus will result in a 25% increased boot time and an 8% increase in the scan time, on average. Although the actual difference might not be significant (in seconds), this could be crucial if heavy, CPU/memory intensive tasks like video editing or compilation are run.

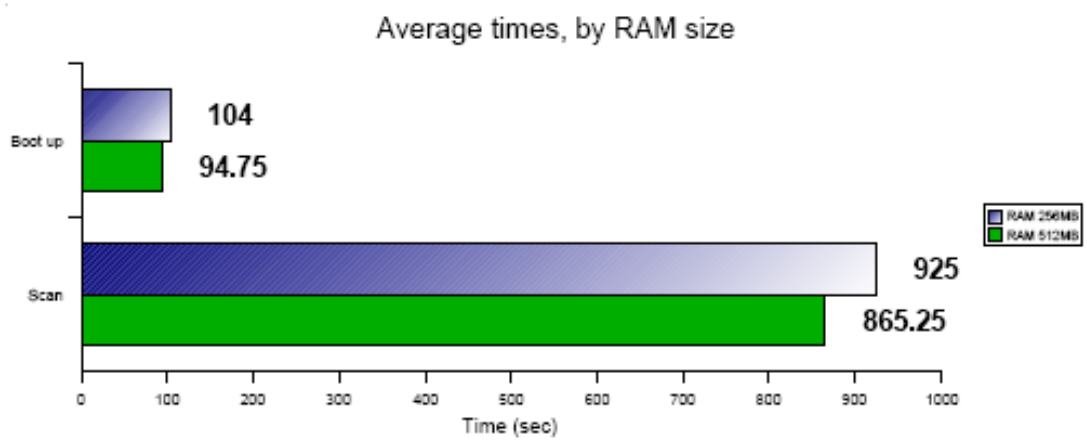
Figure 16: Average times, by Anti-virus



4.3.2 RAM size

Having 256MB of RAM will result in a 10% increase in the boot up time compared to 512MB RAM, and a 7% increase in the scan time.

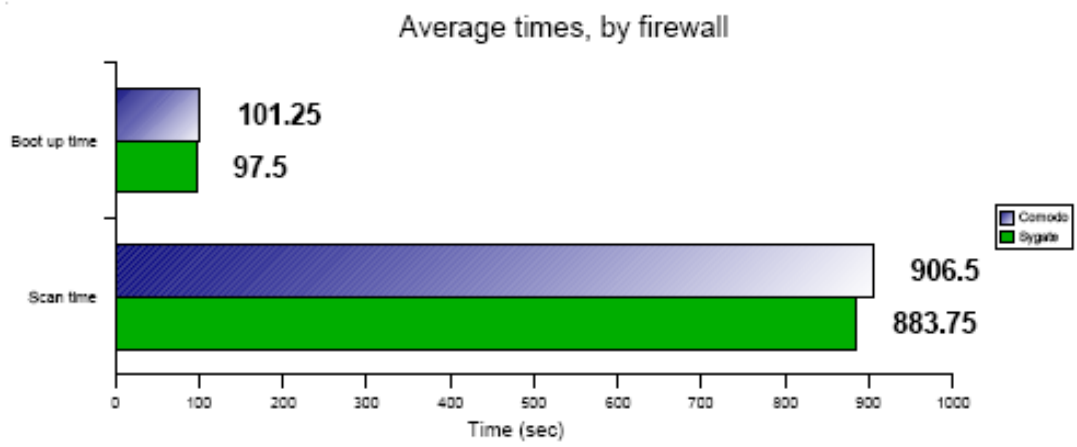
Figure 17: Average times, by RAM size



4.3.3 Firewall

As expected, the choice of a firewall makes the smallest difference. Using Comodo firewall will result in only 4% and 2.5% increases, respectively.

Figure 18: Average times, by firewall



4.4 Miscellaneous facts

Finally, some miscellaneous facts: During the experiment, the AVG anti-virus ran with 4 processes at 3.3MB memory. The McAfee anti-virus ran with 10 processes , at 32.6MB memory. The Sygate firewall ran with 2 processes, at 6.5MB memory. The Comodo firewall also used 2 processes, at 14.3MB. In total, the operating system with the AVG + firewall configuration had 25 processes running. With the McAfee anti-virus, the number of real-time processes was 31. The combined memory usage of the system (other processes also influenced) was as low as 103MB for AVG + Sygate combination and 188MB for McAfee + Comodo. Furthermore, the usage of the AVG anti-virus did not result in frequent CPU spikes, whereas the McAfee product manifested in 15-20% CPU spikes every 6-7 seconds.

5 Conclusions

The most crucial factor for the OS performance is the anti-virus, with as much as 25% impact on the boot up time and 8% impact on the scan time. RAM size is important, but it can be compensated for by the "light" choice of an anti-virus product. The optimal configuration was AVG + Comodo with 512MB RAM, resulting in the shortest boot up and scan times. Overall, Sygate firewall had a slightly better performance over Comodo, especially for the 256MB setting.