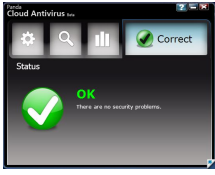


16 May 2009

By: Ionut Ilascu, Editor, Software Reviews



Panda Cloud Antivirus screen

## [The Insides of Panda Cloud Antivirus](#)

*An exclusive interview with Panda's Senior Research Advisor, Pedro Bustamante*

Not long ago Panda [released](#) its Panda Cloud Antivirus product, thus marking a shift in the paradigm of malware detection and removal. Even if in beta, it quickly gained popularity among users thanks to its connection to Collective Intelligence, a huge database of malware signature files. Based on the concept of cloud computing, this innovative approach yields remarkable results in both efficiency and impact on resource consumption.

The application is intelligent and has been designed to act only when necessary taking fast action in detecting and eliminating malware on your machine without affecting the performance of the system. However, after [testing](#) the beta version of this new antivirus model, we were still left with a lot of questions about its inner-workings, as well as Panda's future plans for it.

In our quest to get some answers, we contacted Mr. Pedro Bustamante, senior research advisor at Panda Security, who was kind enough to provide us with some insight into the technology. We invite you to read on as he details the processes behind the first truly cloud-based antivirus solution.

1. The concept of cloud computing has been out there for a while now. How long ago did you get the idea of a "cloud antivirus," and how long did the development for the "[Collective Intelligence](#)" platform and antivirus client last? ([read answer...](#))
2. Panda Cloud Antivirus is a great solution for low-end PCs thanks to its modest use of system resources. Its connection to the Collective Intelligence makes it one of the best protection agents on the market. What would happen if more and more Panda users switched from the paid product that collects the malware from individual PCs and sends it to the cloud, to Panda Cloud Antivirus? How would this affect the cloud and the files it gets? Would this diminish malware detection? ([read answer...](#))
3. What are the sources that provide the "Collective Intelligence" with the information it requires? Do the users of the other Panda products contribute to the cloud? Are files coming from the industry's malware sample exchange channels included? ([read answer...](#))
4. Since it is not an antivirus in itself but more of an agent, so it does not have an AV engine, can Panda Cloud Antivirus coexist with full-fledged protection suites such as Panda Internet Security, Norton AV, Kaspersky, BitDefender or NOD32? ([read answer...](#))
5. Given that Panda Cloud Antivirus offers the best security when connected to the Internet, what would be the percentage drop in protection when working only with the local cache? ([read answer...](#))
6. From our tests, the product scored great when it came to sparing system resources (an average of less than 20MB on Vista). Do you plan on reducing resource usage even more? ([read answer...](#))
7. On the Panda Cloud Antivirus blog, you mention that when there is no Internet connection available, the antivirus will use a locally stored cache of the Collective

Intelligence. Since it's obvious that this cache gets updated regularly, isn't this similar to the signatures update of a classic antivirus program? ([read answer...](#))

8. You also mention on the official blog that, "Non-PE files such as pictures, documents, etc are not scanned from the cloud." Could you elaborate on that? What exactly happens with the latter? For example, maliciously-crafted PDF files, which are a rather common occurrence. ([read answer...](#))

9. It is our understanding that all Panda Cloud Antivirus users are protected from a malicious file within 6 minutes of it being tagged as one. Can you explain for our users in more detail how that file gets to initially be tagged as malware? ([read answer...](#))

10. One statement in your online documentation caught our attention. It reads as follows: "Once the Beta period has finished, the detection capacity of the antivirus will reduce considerably as it will no longer have access to our Collective Intelligence servers. From then on, you can get the free version and benefit from special conditions when you buy the full service." Does that mean that users of the free version will not benefit from the Collective Intelligence unless they pay? Furthermore, on the project's blog, you say that once out of beta Panda Cloud Antivirus will be 100% free. Are companies included as well? ([read answer...](#))

11. Panda Cloud Antivirus does great at detecting malware on relatively clean computers, but how does it behave on already heavily infected machines? And, more importantly, what are its capabilities of terminating malware processes and eliminating them? ([read answer...](#))

12. Is there a deadline for the final release? ([read answer...](#))

13. Do you think you've set a new trend for security applications and that other antivirus companies will bring their own clouds to the user? ([read answer...](#))

14. What do you think of the avalanche of security vulnerabilities in AV developer's websites discovered by grey-hat hackers? Does this affect the image of AV companies? Should this affect users' trust in the security products? ([read answer...](#))

Softpedia: The concept of cloud computing has been out there for a while now. How long ago did you get the idea of a "cloud antivirus," and how long did the development for the "[Collective Intelligence](#)" platform and antivirus client last?

Pedro Bustamante: We initially started working on the idea back in 2006, when we realized that, in order to keep up with the incremental growth of new malicious code, we would need to hire thousands of engineers at PandaLabs to keep our customer well protected. That's when we realized that automation was the only way to deal with unlimited numbers of malware in a faster, more efficient way. In 2007 we published a White Paper called "[From Traditional AV to Collective Intelligence](#)" where we hinted at the fact that we were already working on Panda Cloud Antivirus, at that time codenamed "NanoAV". Since then we've released a few products that took advantage of Collective Intelligence in order to populate the database and to test the system in real scenarios, starting with NanoScan in 2007, ActiveScan 2.0, our consumer products and finally culminating with Panda Cloud Antivirus. ([back to top](#))

Softpedia: Panda Cloud Antivirus is a great solution for low-end PCs thanks to its modest use of system resources. Its connection to the Collective Intelligence makes it one of the best protection agents on the market. What would happen if more and more Panda users

switched from the paid product that collects the malware from individual PCs and sends it to the cloud, to Panda Cloud Antivirus? How would this affect the cloud and the files it gets? Would this diminish malware detection?

Pedro Bustamante: Actually both products work very similar in that respect. Both incorporate from-the-cloud scanning and both send statistics and files to be analyzed remotely. So in that respect the Collective Intelligence and its malware detection capabilities would not be affected at all. The main differences are that Cloud Antivirus has a new agent architecture whose on-access is connected in real-time against Collective Intelligence, while the Panda paid products are complete security suites which include firewall, anti-spam, parental control, identity protection filters, anti-phishing, online and local backup, tuneup, technical support, etc. ([back to top](#))

Softpedia:&nbsp;What are the sources that provide the "Collective Intelligence" with the information it requires? Do the users of the other Panda products contribute to the cloud? Are files coming from the industry's malware sample exchange channels included?

Pedro Bustamante: There's a large amount of sources of information. As you said, pretty much all Panda products contribute, as well as malware samples exchanges from within the industry. But there are also other channels such as CERTs, online scanners such as VirusTotal, customer submissions, honeypots and honeymonkeys we've deployed in a few continents, and a large etc. Of course each source is weighed accordingly, and seeing a piece of code running on a real system is considered the most important for the prevalence algorithms. The community and telemetry aspect of Collective Intelligence is one of the most important parts of the system. ([back to top](#))

Softpedia:&nbsp;Since it is not an antivirus in itself but more of an agent, so it does not have an AV engine, can Panda Cloud Antivirus coexist with full-fledged protection suites such as Panda Internet Security, Norton AV, Kaspersky, BitDefender or NOD32?

Pedro Bustamante: Actually the agent does have a signature + heuristic engine which is optimized to work in off-line mode. That's one of the reasons for the performance optimization we've been able to do. But its not only the engine that makes AVs incompatible with each other. As AVs are some of the most complex applications there exist due to the interactivity they need to have with the host Operating System, the incompatibilities come from the hooks and interceptions they need to perform all over the system. Therefore even though it's a different type of AV, it's still an AV and therefore cannot be run alongside other vendors' AVs. ([back to top](#))

Softpedia:&nbsp;Given that Panda Cloud Antivirus offers the best security when connected to the Internet, what would be the percentage drop in protection when working only with the local cache?

Pedro Bustamante: One of the philosophies of the new protection model we've designed is that AVs don't need to detect every piece of malware that has ever existed or will exist, which is the traditional signature model. Basically what we're saying is that, if we have "x" millions of users in the Collective Intelligence community, what Panda Cloud Antivirus really needs to protect against is whatever malware is circulating amongst those users, and protect them even while off-line. For the rest of malware, you can detect it while connected to the Collective Intelligence servers.

In other words, why do you want your AV to protect you against malware which you'll never even see or come across? Why do you need your AV to waste resources trying to detect

some 5 year old malware, 5 day old malware that's already dead or even malware that is affecting Norton users but not Panda users?

This is where the whole "community" aspect of Panda Cloud Antivirus steps in. Whatever Collective Intelligence "sees" out there as circulating in the wild, it creates a small cache version of signatures which detects and disinfects that subset of malware and synchronizes it in every agent for off-line operation. Even while off-line, Panda Cloud Antivirus will protect against all malware that is circulating, against all malware that is "important for you".

Another way of putting it is that this AV has been designed for real people, real users, not for testers and evaluators which judge how good or how bad an AV is based on lab isolated tests of millions of samples which have not seen the light of day in months or even years. Unfortunately the multi-billion AV industry is very influenced (and therefore limited) by what magazine and independent comparatives publish, even though most testing methodologies in existence today still do not try to reflect the real life situations of end users. We're very hopeful that the work of AMTSO is going to help a lot in improving testing methodologies and bring them closer to reflect real life scenarios. ([back to top](#))

Softpedia: From our tests, the product scored great when it came to sparing system resources (an average of less than 20MB on Vista). Do you plan on reducing resource usage even more?

Pedro Bustamante: Yes, that's the objective, although it'll become harder as we try to improve the product. Keep in mind it's still in beta and we're still adding some other technologies to it. Our internal development target is to stay below 20MB. That's what we've defined as our absolute top when running silently in the background.

However the most important metric we're concentrating on is Performance Impact, which is a measurement of how much impact a specific application has on a clean machine, based on measuring different operations such as file copy, decompression & compression, browsing, opening applications such as Word, etc. There's a few standards for this, such as the WorldBench Index. The AV industry average in this metric is around 10% performance impact on a clean system. Our target is to stay below 5%. In the case of Windows XP we've managed to bring it down to a little over 3%, so that's already a huge improvement. ([back to top](#))

Softpedia: On the Panda Cloud Antivirus blog, you mention that when there is no Internet connection available, the antivirus will use a locally stored cache of the Collective Intelligence. Since it's obvious that this cache gets updated regularly, isn't this similar to the signatures update of a classic antivirus program? Please elaborate.

Pedro Bustamante: Yes, as mentioned above, Panda Cloud Antivirus keeps a local cache of Collective Intelligence which gets updated as needed. However, it's not quite the same as the traditional signature updates which are always incremental (always adding signatures, not taking them out). The local cache of Panda Cloud Antivirus is a "moving target" of what the community "sees" out there circulating in the wild. The fact that we're not relying on signature updates as the main protection model anymore, relieves us from having to push down very frequent updates to the client, with the bandwidth consumption and CPU/RAM consumption for patching and loading the signature file which that process ends up eating.

The local cache also includes other types of generic signatures, generic disinfection routines and non-PE signatures. These are used mostly for off-line operation and for certain type of malware. The local cache contains less than 10% of the full knowledge of Collective

Intelligence. We believe that as malware becomes more and more dynamic and the number of total malware continues growing exponentially, this % will be reduced over time. ([back to top](#))

Softpedia: You also mention on the official blog that, "Non-PE files such as pictures, documents, etc are not scanned from the cloud." Could you elaborate on that? What exactly happens with the latter? For example, maliciously-crafted PDF files, which are a rather common occurrence.

Pedro Bustamante: Basically detection of scripts, macro viruses, PDF files, and other types of files such as file-infecting viruses is included within the local cache signatures of Cloud Antivirus. We've designed Collective Intelligence to be able to automatically analyze and protect against PE files, which is really the vast majority of the malware problem. Developing something similar for non-PE files would have been too costly for the small benefit that we would have gained from it.

However for maliciously crafted PDFs, DOCs, PPTs, and similar files we do have a very efficient technology, called "Kernel Rules Engine" which is able to generically detect 100% of these types of exploits without any signatures. This is one of the improvements we'll integrate into Panda Cloud Antivirus in the short future. More info about this technology [here](#). ([back to top](#))

Softpedia: It is our understanding that all Panda Cloud Antivirus users are protected from a malicious file within 6 minutes of it being tagged as one. Can you explain for our users in more detail how that file gets to initially be tagged as malware?

Pedro Bustamante: There are a variety of processes running behind the scenes in Collective Intelligence. They can be broken down in four main phases: analysis, categorization, remediation and telemetry.

Behind the initial analysis phase there's a bunch of technologies that are used to extract all type of information from each file, both from static analysis (such as packer information, API calls, functions, multi-scanners, etc.) and dynamic analysis (running in real machines, recording malware actions, dumping memory, etc.). All this information is then processed in the categorization phase, where it is correlated against the entire database of Collective Intelligence files using different techniques, such as graph theory algorithms, grouping algorithms, metaheuristics, rule driven classification and identification, and many more techniques which are too resource intensive that can only be run in a server-farm environment such as Collective Intelligence and not on end user PCs.

The remediation phase consists basically of creating detection signatures and disinfection routines which are served to agents both through cloud-requests and, in case the telemetry and community statistical phase so decides, pushed down to local caches of Panda Cloud Antivirus agents. ([back to top](#))

Softpedia: One statement in your online documentation caught our attention. It reads as follows: "Once the Beta period has finished, the detection capacity of the antivirus will reduce considerably as it will no longer have access to our Collective Intelligence servers. From then on, you can get the free version and benefit from special conditions when you buy the full service." Does that mean that users of the free version will not benefit from the Collective Intelligence unless they pay? Furthermore, on the project's blog, you say that once out of beta Panda Cloud Antivirus will be 100% free. Are companies included as well?

Pedro Bustamante: This is an unfortunate grammar mistake which has caused many users to doubt whether Panda Cloud Antivirus will continue being free after beta or not. Basically what it tries to say is that the "beta version" (not the release version) will stop connecting to Collective Intelligence servers or, in other words, that the beta version will expire. This of course will only happen when the final version 1.0 comes out and when that happens, version 1.0 of Panda Cloud Antivirus will continue being free and will of course benefit from Collective Intelligence connectivity just as the beta version does.

In short, Panda Cloud Antivirus will continue being free and with full functionality & connectivity to Collective Intelligence after the beta version expires. Regarding companies, this is still up in the air but most likely it will be free for home and personal user only. ([back to top](#))

Softpedia: Panda Cloud Antivirus does great at detecting malware on relatively clean computers, but how does it behave on already heavily infected machines? And, more importantly, what are its capabilities of terminating malware processes and eliminating them?

Pedro Bustamante: In short, not as good as it should be, and definitely not as good as when we release version 1.0. Keep in mind that we're still incorporating certain detection and protection techniques into the product such as better anti-stealth, self-protection mechanisms and detection of in-memory patching techniques. However even without all these improvements some independent tests have already revealed that Panda Cloud Antivirus is able to disinfect an infested system in some cases better than full paid products out there, so we're very confident that, once v1.0 is finished, it'll be some of the absolute best, if not the best, AV tools out there. ([back to top](#))

Softpedia: Is there a deadline for the final release?

Pedro Bustamante: Yes, we are hoping to release v1.0 before the end of the year... "2009 that is" - that's an internal joke at Dev :)

But of course that will all depend on whether we are able to finish integrating all the technologies and they work as they should during the testing phase. In this regard we're very happy as we're getting a lot of help and feedback from our beta users, who are doing an excellent job of providing detailed information about things which need improvement. ([back to top](#))

Softpedia: Do you think you've set a new trend for security applications and that other antivirus companies will bring their own clouds to the user?

Pedro Bustamante: Absolutely, we have a history of doing precisely that. In 1998 we started pushing daily signature updates and hooking the Winsock to scan all TCP/IP traffic transparently at the desktop. A few years later the rest of AVs were doing that same thing. In 2004 we introduced a behavioral analysis and blocking engine, called TruPrevent, into desktop protection. To date there's still quite some mainstream AVs that are still developing their own behavioral engine (or even buying the technology instead of developing it).

In 2007 we released the first application of cloud-scanning technology and even today only a handful of AV manufacturers have announced they're working on something similar. Of course only time will tell, but I do believe it is the natural evolution for an AV. The benefits of cloud-computing are huge and can be very effective for AV protection. However it does require a big change in mentality as mentioned in some responses above, like prioritizing

some samples over others, and definitely a very large investment in infrastructure. ([back to top](#))

Softpedia: What do you think of the avalanche of security vulnerabilities in AV developers' websites discovered by grey-hat hackers? Does this affect the image of AV companies? Should this affect users' trust in the security products?

Pedro Bustamante: There still exist a lot of publicity-driven hackers, and security vendors are a juicy target for them. Be it website XSS hacks or simple file format parsing and scan evading Proof of Concept code (which by the way are never used by real malware writers), this is something we have to live with. Unfortunately the media gives it much more importance than it really has. There are hundreds if not thousands of more dangerous hacks & PoCs happening every day in mainstream websites and applications, yet yellow journalism seems to dictate that it is only newsworthy when a security vendor is involved. It's a very strange thing, but there's a lot of emotions and love/hate relationships with AV programs and AV manufacturers. You don't see that in any other software industry. ([back to top](#))