Kaspersky®
Security Network

Kaspersky Security Network: Big Data-Powered Security

Cloud intelligence for protection against complex threats

The number of cyberattacks globally increases every day. This has a significant impact on business, often resulting in data theft or loss of important information, threatening the business processes of companies of every type and size, from local start-ups to multinational industry leaders. And it's not just about the number of attacks – new generations of malware appear all the time, many using new, sophisticated techniques to bypass existing security solutions.

In this constantly shifting environment, protection is only as effective as a vendor's ability to closely monitor the threat landscape and distill data into actionable intelligence for its customers. To achieve this, security solutions must apply a cloud approach that combines the widest possible scope of threat data collection with the most intelligent data processing technologies.

Kaspersky Security Network (KSN) is a complex distributed infrastructure that integrates cloud-based technologies into Kaspersky Lab's products. KSN automatically analyzes cybersecurity-related data streams from millions of voluntary participants around the world, providing the fastest detection of advanced and previously unknown malware.

The key components of KSN success:

- The acquisition of global malware detection statistics along with real-time data on suspicious activities
- A powerful combination of Big Data analysis, machine learning and human expertise to process the collected information
- Fast delivery of threat intelligence to customers – leaving resource-intensive computation in the cloud.

The hardest part is sorting and analyzing the data – the volumes are so huge that we use data science-based automation, including several methods of artificial intelligence, to process it. The human element remains an important advantage of this system, because only human intuition and experience can help machines to cope with the complex and often highly imaginative creations of malware authors. Kaspersky Lab's experts have real-time access to all the information being gathered, enabling them to gain valuable new insights into threats, apply knowledge to investigations and develop new, proactive detection technologies.

This approach offers numerous benefits for customers, including:

- Best detection of advanced and previously unknown malware
- Reduced detection errors (false positives)
- Significant reduction in response times to new threats – traditional signature-based responses can take hours; KSN takes about 40 seconds.

KSN privacy protection

KSN processes data from a user's device only after a user accepts the data processing agreement which describes the full scope of data collected and the purposes of its processing. The information used in data processing is maximally anonymized, data is stored and processed in accordance with strict security policies.

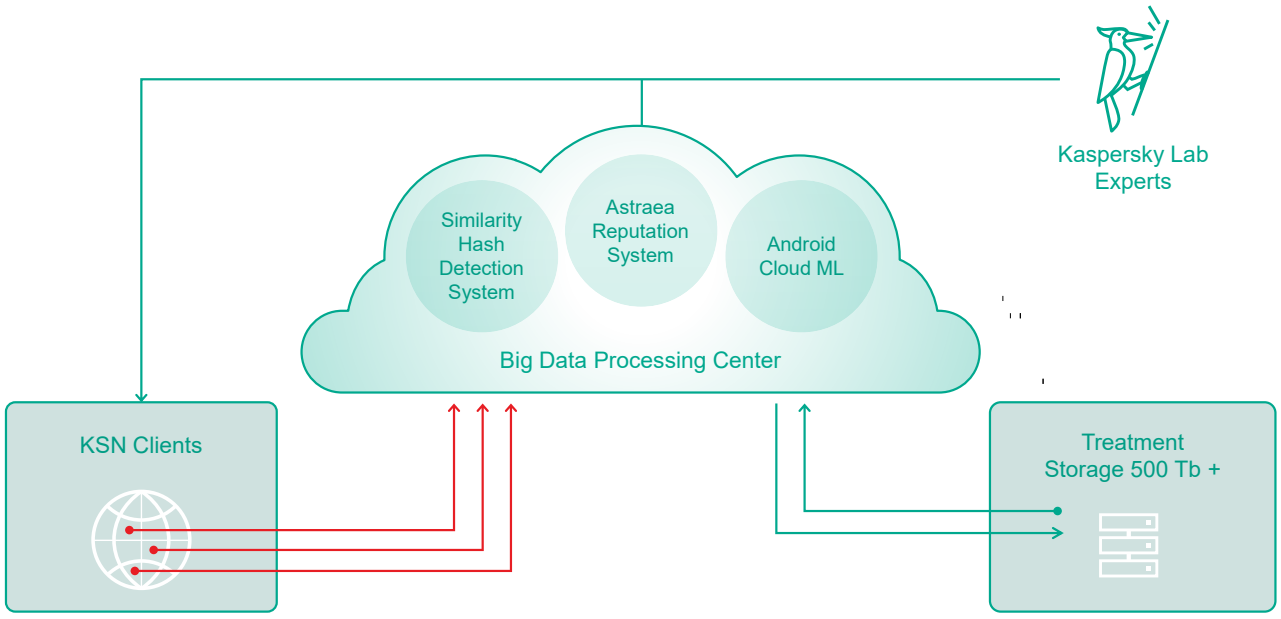

Fig.1 – Scheme of KSN elements

Astraea: Reputation system based on Big Data

Hundreds of millions of different records come in to KSN, a massive volume of data often reaching hundreds of gigabytes daily. This anonymized data is compressed and stored for future use; even after compression, it still requires terabytes of storage.

One of the systems Kaspersky Security Network uses to process this enormous data stream is called Astraea. Every day, Astraea processes information on millions of objects, sorting and analyzing it. Once sorted, it rates each object (only the metadata of the object used for the analysis, not its contents).

Every suspicious event received by the system is evaluated according to importance and potential danger using multiple criteria. Following this analysis, the object's reputation is calculated, and global statistics about it are requested. What else can the collective intelligence tell us about it? Perhaps its reputation is even worse than it initially seems? Or is it a false alarm? This querying against other information allows the system to fine-tune verdicts and reduce the probability of false positives.

When an object's accumulated statistics confirm that it's malicious, secure or has an unknown status, that information is made available across all supported Kaspersky Lab products where users have enabled Kaspersky Security Network – without any human intervention. Similarly, when malicious web resources are processed, users automatically receive a warning of the danger when they try to access it.

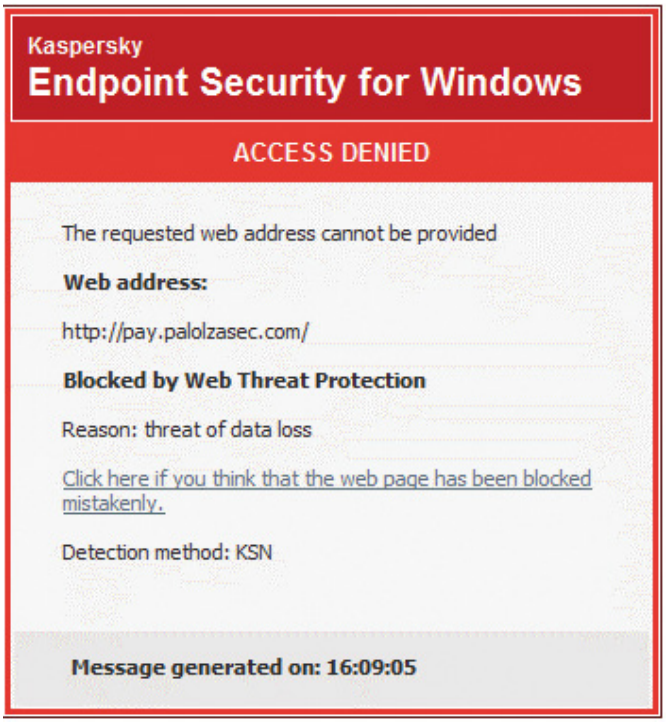

Fig. 2 – Dangerous site alert

Astraea is a highly intelligent system that constantly learns to deal with the rapidly changing threat landscape. However, for all the advantages of automation, protection without people is impossible, because the system has to understand to withstand the inventive detection evasion techniques of human cybercriminals. That is why KSN (like other Kaspersky Lab products and systems) uses the HuMachine principle – combining machine power and human expertise.

How does this work? If Astraea can't determine the level of threat posed by an object, the data is sent to human experts who conduct additional in-depth analysis before adding the data to KSN for instant detection through the cloud. At the same time, heuristic detection models can be adjusted to detect many different malware specimens based on the similarity of indicators..

Machine learning in KSN

Together with the Astraea reputation system, two cloud ML-based technologies are available through KSN to proactively detect unknown threats. In both cases, ML models are created and trained in-lab, while an endpoint or mobile product just applies them for fast, real-time detection.

Similarity Hash Detection

It used to be that hashes were used to create malware "footprints" sensitive to every small change in a file. This drawback was exploited by malware writers through obfuscation techniques like server-side polymorphism: minor changes in malware took it off the radar. KL products address this threat with several advanced methods to detect malware variations. Similarity hashes (SH) is one of them.

The cloud component of the Similarity Hash Detection system collects multiple file features from different sources including in-lab automatic systems of malware processing. Then a machine learning approach is used to find the features common to the whole group of similar malicious files. Based on these features, similarity hashes are calculated and published through KSN.

The endpoint product extracts a file's features at the endpoint, calculates its SH and checks it through both local and cloud SH databases. Even if a new malware file is somewhat different from a known one, its SH will be similar or identical to the SH already known to KSN. This approach allows Kaspersky Lab products to detect whole families of quickly changing polymorphic malware.

Android Cloud ML

In our ML-based system of mobile threats detection, the predictive model takes the form of a Decision Tree Ensemble. Every non-leaf node of a tree contains a question about the features of the file being examined, while the leaf nodes contain decisions on the file; decisions of multiple trees are averaged in an algorithm-specific way to provide the final classification of the file: malicious or not.

This type of powerful ML model trained on millions of samples can detect malware with high accuracy but it requires lots of resources to run which would be hard to provide on a simple mobile device. This is where we benefit from the cloud approach. First, the agent on a user's device collects multiple features of an Android application – its entry points, permissions, etc. – to get the most accurate description of the app (no sensitive user data is collected). This data vector is sent to the KSN cloud where it's run through the Android Cloud ML model, and its classification decision is immediately sent back to the mobile device.

Expertise for the real world

Threat intelligence delivery from KSN is fast – it's measured in seconds. This maintains a consistently high level of protection against real-world, real-time cyberthreats. In the event of a mass attack, when information about the malware has already reached KSN servers but has not yet been delivered to end-users in the form of detection records, the system will provide it immediately in response to a user request.

Of course, this requires some bandwidth, and Internet traffic may be limited. KSN can use local caching servers installed within the local network, which helps to reduce the load on the Internet connection. If a user has switched KSN off, they only receive information about new malware following an update – until then, they continue to be protected by other proactive mechanisms.

KPSN

While all information processed by Kaspersky Security Network is completely anonymized and disassociated from source, Kaspersky Lab recognizes that some organizations – for compliance or company policy reasons – require absolute data lockdown. This has traditionally meant that enterprises can't avail of cloud-based security services.

For these customers, Kaspersky Lab has developed a standalone product: **Kaspersky Private Security Network**, which allows enterprises to take advantage of most of the benefits of global cloud-based threat intelligence without releasing any data whatsoever beyond their controlled perimeter. It's a company's personal, local and completely private version of Kaspersky Security Network.

KPSN can be installed on a special local server, providing adaptive protection to all connected devices. KPSN does not allow access to the Internet – in an especially strict environment, updates can be done manually using secure portable media. Either way, the provision of an inbound data stream greatly boosts reaction times to constantly shifting threats:

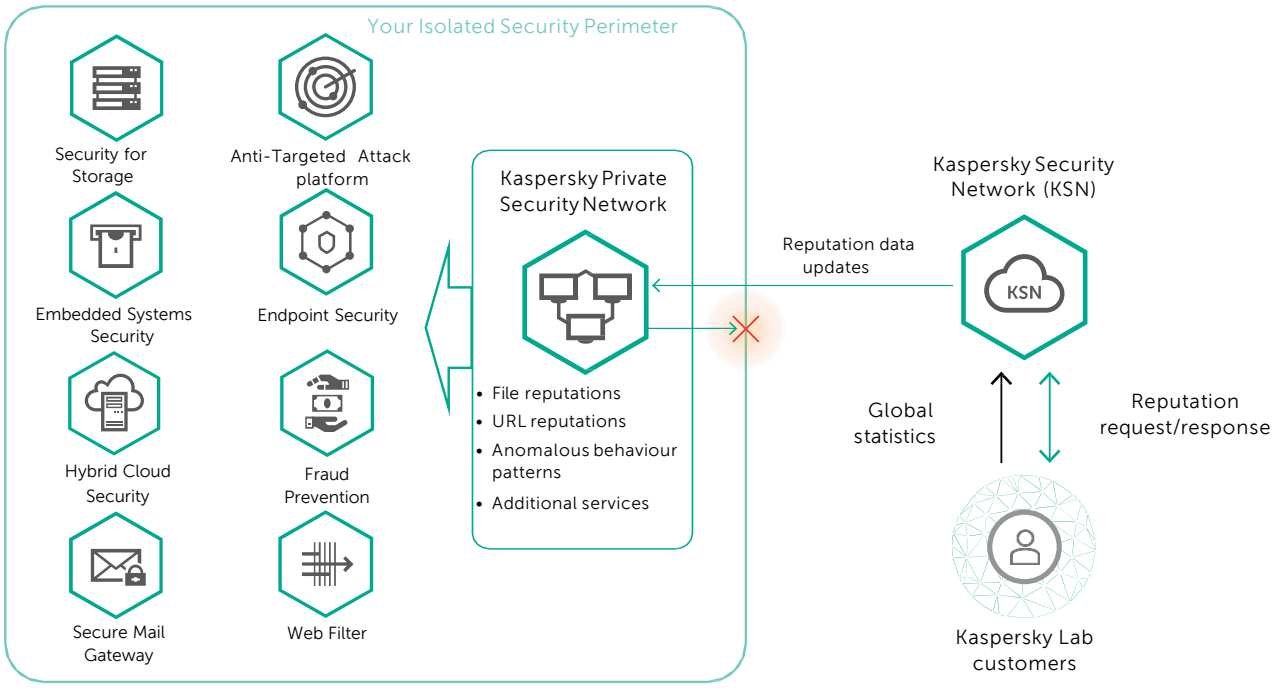

Fig.3 – Scheme of KPSN infrastructure in secured perimeter

Conclusion

The need for immediate protection against new threats is obvious, even to those who aren't directly involved in information security. Even without Kaspersky Security Network enabled, the multiple layers of protection technologies in our solutions provide effective security for all users.

What KSN's instant, cloud-based protection adds to the mix is support for additional, important security mechanisms that minimize false positives while increasing the quality of detection using real-time, additional data on threats, authorized applications and other relevant information.

Complex and targeted threats tend to cause significantly greater damage than mass malware. KSN and KPSN protect your business against these threats, mitigating attack risks and reducing costs while ensuring data security and business continuity.

All about Internet security: **www.securelist.com**
Find a partner near you: **www.kaspersky.com/buyoffline**

www.kaspersky.com
#truecybersecurity