Google Cloud | VIRUSTOTAL

VirusTotal File Feed

Continuous, retroactive enrichment to unearth historically unnoticed intrusions that persist in your environment

Organizations starting their journey in the threat intelligence space enrich their telemetry by performing a VirusTotal API lookup for every observable found in their alerts. They do so in a fire and forget fashion, where historical events get buried in some SIEM or TIP, neglected and completely inactionable. The file feed allows you to sync historical events with VirusTotal's brain in order to unearth undetected threats that originally flew under the radar. This is how mature security teams make breaches insignificant. Additionally, the feed incorporates rich IoC relationships that can be fed into your perimeter defenses in order to implement a preventative strategy, or to simply enjoy the best of VirusTotal on-premise.

Mitigate the impact of stealthy breaches

Perform automated, continuous, retroactive enrichment of historical observables archived in your SIEM, flag suspicious historical patterns as soon as new findings are uncovered by the community. Track threat actors and proactively uncover intrusion toolkits still unknown to the industry.

Preventatively neutralize threats

Use data analytics to automatically digest the feed into hunting and preventative IoCs. Feed these into your existing tech stack and take your security investments to the next level by proactively blocking threats before they hit you.

Comply with regulatory requirements

Work with VirusTotal's live malware flux on-premise or recreate our dataset in your own infrastructure to conduct highly sensitive investigations inline with your secops and policy requirements. Benefit from VirusTotal's knowledge in air-gapped environments.

Boost your defense-in-depth strategy

Complement your existing security solutions, cover their blindspots and add an additional layer of defense based on crowdsourced context.

In a nutshell ➤➤ ➤

Pubsub-like live stream of all analysis reports & metadata generated by VirusTotal for uploaded files, along with a link to download these if desired.

Massive ›› 1M+ file analyses per day, 500K+ detected 1+ AV vendors

Diverse ›› Contributions by 2M+ monthly users from 232 countries

Rich ›› Industry threat reputation, {YARA, SIGMA, IDS} detection crowdsourcing, static dissection, provenance details, submission and in-the-wild metadata, etc.

Real-time ›› Ingest new threat details on-the-fly as the community discovers and uploads them to VirusTotal

Actionable ›› Related IoCs and TTPs thanks to sandboxing and threat graph interlinking

Google Cloud | VIRUSTOTAL

File feed use cases

Your path to stronger, more proactive cybersecurity operations starts here

On-premise VirusTotal dataset replica

Organizations operating in highly regulated segments or air-gapped environments are often unable to perform API lookups for observables encountered in their investigations. The file feed allows them to ingest absolutely all file reports live, as the analyses conclude. These reports can be inserted in a local database in order to then perform lookups against such replica.

On-premise YARA Livehunt

Intelligence agencies and critical sectors are often subjected to regulations or secops practices that prevent them from storing investigative artifacts such as YARA rules in 3rd-party vendor platforms. In those cases they can license the file feed to download absolutely every single file uploaded to VirusTotal, real-time, and match them with YARA locally in their own on-premise infrastructure.

Automated, continuous, retroactive enrichment

Performing an initial API lookup whenever a new observable is encountered is highly suboptimal, as the threat could still be unknown to the industry. Performing periodic lookups of historical events is not the answer either, as your visibility would only be as good as your lookup frequency. The file feed acts as a pubsub stream allowing you to enrich, real-time, historical sightings. Thanks to this you may surface originally undetected breaches as soon as the security community uncovers new threats or has new findings for them in terms of network infrastructure.

Automated preventative & hunting IoC generation

Dump the incoming stream of file metadata and related indicators into a data analytics solution such as Splunk or Google Cloud Bigquery. Perform regexes, group by's and sortings in order to instantly spot threat campaign/malware toolkit commonalities that can be used to feed your existing security technologies. Prevent attack variants and threats reusing malicious network infrastructure (CnCs, dropzones, etc.). No need to store massive amounts of reports, a continuously updated dump of the last week or last month of data will suffice.

Non-AV competitive machine learning

Antivirus vendor verdicts should not be used to train machine learning models that will then compete against the vendors themselves. This said, there are numerous other applications of ML that you can indeed implement, the simplest one would be to train a Naïve Bayes spam filter with emails uploaded to VirusTotal that contained malicious attachments. Another example is to improve your defenses against malicious documents by training a model with malware macros.

Custom filtering & matching beyond YARA

YARA is the de-facto standard for campaign monitoring and threat actor tracking, in VirusTotal you can not only match file binary contents, *but also metadata and dynamic analysis properties.* This said, sometimes these pattern matching capabilities are not enough and you may want to apply transformations/functions or to join VirusTotal's knowledge with other licensed datasets in order to perform advanced matching. The file feed allows you to solve this by downloading the files themselves and their reports.

Malware execution in specialised environments

Execute the VirusTotal live file flux in your own sandboxes or analysis pipelines in order to uncover zero day malware or to assess the impact of a given threat when detonated in your organization's systems.

Stack validation and breach & attack simulation

Continuously test your security infrastructure with a live feed of malicious files and indicators so as to identify blind spots, misconfigurations and unnoticed settings changes that can negatively impact your security posture.

Surface unknown tools used by your adversaries

File reports published in the feed contain submission metadata: dates, file names, submission country and, most importantly, a ciphered submitter identifier. By creating a local mapping between this anonymized identifier and their historical uploads you may retroactively uncover previously unknown malware uploaded to VirusTotal by a potential threat actor.

Clustering & malware family identification

Apply your own custom logic to the submitted files or their report properties in order to create groupings that surface polymorphic behaviours. Intercept new malware campaigns and implement appropriate defenses against these. This does not require you to store years of file feed data, but rather a continuously updated dump of the last week or month.