

KAVYA SRINIVAS MURTHY

CENSYS

- * It's an all-new Hacker's Search Engine similar to Shodan, which is designed specifically to locate any devices that have been carelessly plugged into the Internet without much attempt at preventing unauthorized access.

- * The leading internet intelligence platform for threat hunting and exposure management

- * Censys empowers security teams with the most comprehensive, accurate and up to date map of the internet to defend attack surfaces and hunt for the threats.

- * Not all the data is created equal-- Censys search empowers governments, enterprises, and researchers with the most comprehensive, accurate, and up to date Map of the Internet for proactive and reactive security analysis at scale.

- * At the end of last month, security researchers from SEC Consult found that the lazy manufacturers of home routers and Internet of Things (IoT) devices have been re-using the same set of hard-coded cryptographic keys, leaving around 3 million of IoT devices open to mass hijacking.

- *But how did the researchers get this number?

→ Researchers uncovered these devices with the help of Censys – a new search engine that daily scans the whole Internet for all the vulnerable devices.

- *Censys Maintains Complete Database of Everything on The Internet.

- * However, Censys employs a more advanced method to find vulnerabilities in the devices and make the Internet a safer place.

- * Censys is a free search engine that was originally released in October 2015, by researchers from the University of Michigan and is powered by the world's biggest search engine Google. However, it emerged as a company in 2017. Censys was developed by Durumeric, graduate student David Adrian, fourth-year undergraduate Ariana Mirian, and Prof. J. Alex Halderman, all of Michigan, along with Prof. Michael Bailey of UIUC. Their research paper on Censys, entitled “A Search Engine Backed by Internet-Wide Scanning,” appeared at the 22nd ACM Conference on Computer and Communications Security (CCS) in October 2015. The paper contains a full description of Censys’s architecture and several use cases.

- * Censys is part of an open-source project that aims at maintaining a "complete database of everything on the Internet," helping researchers and companies unearth Online security mishaps and vulnerabilities in products and services.

- * Censys, a cloud-based service that not only maintains an up-to-date snapshot of the hosts and services running across the public IPv4 address space, but also exposes this data through a search engine and API.

- * In contrast to existing scanning tools, which have primarily focused on performing host discovery, Censys immediately produces results based on full protocol handshakes, facilitates a community-driven approach to characterizing the exploding number of embedded devices and vulnerabilities on the Internet, and requires little or no user preparation.

- * Censys continually scans the public address space across a range of important ports and protocols. It validates this data and performs application-layer handshakes using a pluggable scanner framework,

which dissects handshakes to produce structured data about each host and protocol. The resulting data is post-processed with an extensible annotation framework that enables researchers to programmatically define additional attributes that identify device models and tag security-relevant properties of each host.

Censys transparently and expose data back to the research community. Censys exposes data to researchers through a public search engine, REST API, publicly accessible tables on Google BigQuery, and downloadable datasets.

The search interface enables researchers to perform full-text searches and query any of the structured fields and tags produced during scanning and post processing (e.g., 443.https.cipher_suite). It supports full-text searches, regular expressions, and numeric ranges, and queries can be combined with Boolean logic.

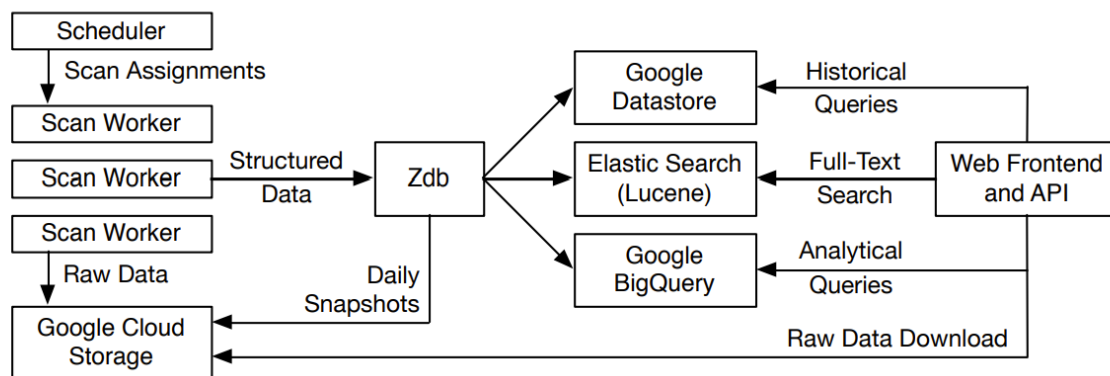
These queries can be run against a current snapshot of publicly accessible IPv4 hosts, Alexa Top 1 Million websites, and known X.509 certificates. After running a query, users can interactively explore the hosts, sites, and certificates that match their query, as well as generate statistical reports suitable for direct use in research.

Advantages of censys:

Censys allows the security community to increase global protocol coverage and provides a tractable solution for understanding the increasing number of embedded devices on the Internet.

Simultaneously, it minimizes redundant scanning by research groups and minimizes the incoming network traffic monitored by network operators.

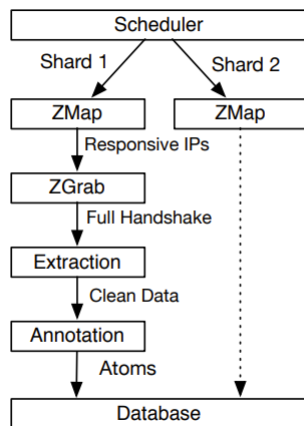
Censys Architecture:



Censys System Architecture — Censys is driven by application scans of the IPv4 address space, which are scheduled onto a pool of scan workers. These workers complete scans, extract valuable fields, and annotate records with additional metadata in order to generate structured data about each host. These records are centrally managed in a custom database engine, ZDb, which maintains the current state of every host. ZDb feeds updated records to a web front-end where researchers can query the data.

Protocol Scanning and Annotation:

Each scan worker uses ZMap to perform host discovery for a shard of the IPv4 address, and completes protocol handshakes using pluggable application scanners. Censys extracts fields of interest and annotates records with additional metadata. The information from a protocol handshake is converted to an atom—a deterministic data structure describing a specific protocol on a host.



How Does Censys Work?

- ❖ Censys collects information on hosts and websites via daily scans of the IPv4 address space – the internet protocol version 4 that routes the majority of the Internet traffic today.
- ❖ This search engine uses two companion tools:
 - 1) ZMap – an open-source network scanner, the hosts found by ZMap seed pluggable application scanners, which perform a follow-up application layer handshake and produce structured JSON data describing a certain aspect of how a host is configured.
 - 2) ZGrab – an application layer scanner, which meets the previous specifications and facilitates the rapid development of new types of scans. At this time, ZGrab supports application handshakes for HTTP, HTTP Proxy, HTTPS, SMTP(S), IMAP(S), POP3(S), FTP, CWMP, SSH, and Modbus, as well as StartTLS, Heartbleed, SSLv3, and specific cipher suite checks.
- ❖ Censys then maintains a database of how hosts and websites are configured, allowing researchers to query the data through a search interface, report builder, and SQL engine.
- ❖ ZMap scans over 4 Billion IP addresses on the Internet and collects new data every day. It also helps determine whether the machines on the internet have security vulnerabilities that should be fixed before being exploited by the hackers.

"We have found everything from ATMs and bank safes to industrial control systems for power plants. It's kind of scary," said Zakir Durumeric, the researcher leading the Censys project at the University of Michigan.

How to Start with Censys Search

Follow the steps below to see the Internet in a whole new way.

1. Sign Up for a Censys Account (www.censys.io)

Before you can start using Censys Search, you need to sign up for a free account.

2. Learn about the Censys Search Language Or optionally use [CensysGPT](#) to translate questions into queries.

3. Start Searching with Queries (Discussed in the later part of document with screenshots).

4. Refine Your Queries as per requirement with values relevant to your investigations

5. Use the Censys API

The Censys Search API allows you to integrate Censys data into your applications, scripts, or security tools.

6. Follow Censys Search Best Practices

To make the most out of Censys Search, follow our best practices:

- Stay informed: Keep up with the latest changes and updates to Censys by regularly checking the official documentation and blog posts.
- Respect privacy and legalities: Be aware of privacy concerns and legal considerations when using Censys Search. Make sure that your use of the tool complies with relevant laws and regulations.
- Experiment and learn: The best way to become proficient with Censys Search is to experiment with different queries and explore its capabilities.
- Collaborate and Share: Share your findings with the community and collaborate with others in the field. Knowledge sharing is vital in cybersecurity.

Search Interface

The primary interface for Censys is a search engine that allows researchers to perform full-text searches and structured queries against the most recent data for IPv4 hosts, the Alexa Top 1 Million websites, and known certificates.

For example, a researcher can find all hosts currently vulnerable to Heartbleed in the United States with the query: `443.https.heartbleed.vulnerable: True AND location.country_code: US`. This query executes in approximately 250 ms and users are presented with the hosts that meet the criteria, along with basic metadata, which includes the breakdown of the top ASes, countries, and tags. Users can view the details of any host, as well as generate statistical reports. Different kinds of searches has been discussed further in the document.

Viewing Individual Records.

Users can view the details any host, certificate, or domain returned by a query. This includes a user-friendly view of how each service is configured, the most recent raw data describing the host, userprovided metadata and tags, and historical scan data. We similarly display geographic location, routing, and WHOIS information.

Below discussion is based on the website censys.io, exploring the search page:

Introduction to Hosts

The Censys Host dataset provides accurate, up-to-date records that reflect the reality of public IPv4 and IPv6 hosts and virtual hosts (i.e., host services reached by name).

Host Records

Host Records are identified by an IP address.

The Censys model of Internet hosts contains more than just service data observed in scanning. Censys also enriches hosts with quality data from third-parties to provide information such as geographical location, network routing information, DNS names, etc.

Note: 1) Virtual hosts are identified by a name, IP address tuple.

2) The difference between a keyword and a text field is that searches on keyword fields will only return exact matches, while searches on text fields will return fuzzy matches.

Host Fields

The Data Definitions page lists every field that can appear in a host or virtual host record.

The screenshot shows the Censys search interface. The browser address bar displays `search.censys.io/search/definitions?resource=hosts`. The page header includes the Censys logo, a search bar with a dropdown menu set to 'Hosts', and a search button. Below the header, there are navigation links: 'Getting Started', 'Search Language', 'Examples', 'Data Definitions' (which is highlighted), and 'Changelog'. A note states: 'This page lists every field whose value can be searched within the Hosts dataset. The difference between a *keyword* and a *text* field is that searches on keyword fields will only return exact matches, while searches on text fields will return fuzzy matches.' Below this note, there is a section titled 'Hosts Categories List' with a 'Collapse All' and 'Expand All' link. The list is organized into four columns. The first column lists categories like 'Host Information', 'Service Information', 'Host DNS', etc. The other three columns list specific fields like 'COAP', 'L2TP', 'PROMETHEUS', etc. A small orange circle with the number '81' is visible in the bottom right corner of the page.

| Category | Field | Field | Field |
|------------------------|---------------|-----------------------|-------------|
| Host Information | COAP | L2TP | PROMETHEUS |
| Service Information | COBALT_STRIKE | LDAP | RDP |
| Host DNS | CWMP | MEMCACHED | REDIS |
| Host Location | DARKCOMET | MMS | ROCKETMQ |
| Host Operating System | DARKGATE | MODBUS | S7 |
| Host Autonomous System | DHCPDISCOVER | MONERO_P2P | SIP |
| LABELS | DNP3 | MONGODB | SKINNY |
| Software | ELASTICSEARCH | MQTT | SMB |
| TLS | ELF_FILE | MSSQL | SMTP |
| HTTP | EPMD | MYSQL | SNMP |
| SSH | ETHEREUM | NTP | SOCKS |
| TELNET | FORTIGATE | OPC_UA | SSDP |
| FTP | FOX | OPENVPN | TEAM_VIEWER |
| DNS | IKE | ORACLE | TPLINK_KASA |
| ACTIVEMQ | IMAP | PC_ANYWHERE | UPNP |
| AMQP | IPMI | PENDING_REMOVAL_SINCE | VNC |
| ANY_CONNECT | IPP | POP3 | X11 |
| BACNET | KRPC | POSTGRES | ZEROMQ |

Top-Level Host Fields

Top-level host fields include information that applies to the host as a whole such as its geographic location, network routing information, DNS names, operating system, and labels, and a repeated record of services observed in scan.

Service Fields

Service records contain identification and metadata fields, labels for easy searching, a protocol-specific sub record with information parsed from scan, TLS fields, and software fields.

Identification and Observation Fields

Identifying service information includes fields like:

- ➔ Service name: The name of the service. Correlates loosely to OSI Layer 7 protocol names. Censys can detect more than 108 services.
- ➔ Port: The port number of the port.
- ➔ Extended service name: The name of the service, including the TLS indicator, if the service negotiated a TLS connection.
- ➔ Transport protocol: The name of the transport protocol. Correlates to OSI Layer 4 protocols (e.g., TCP, UDP, QUIC)
- ➔ Truncated: Whether the service data was truncated because it is a suspected low-value pseudo-service on a super host.

Observation information are fields that provide data about Censys' discovery and observation of the service:

- ➔ Perspective ID: The name of the ISP that Censys peered with when it observed the service as represented.
- ➔ Source IP: The IP address of the Censys scanner when it observed the service represented.
- ➔ Discovery method: The name of the method that led to the discovery of the service by a Censys scanner.

Service-Name-Specific Fields

The data Censys observes about an HTTP service is very different from that of an SSH service, so the parsed data from each scan is searchable within a record that matches the service name.

The fields found in each service-name-specific field reflect the details of that protocol.

Ref link : https://support.censys.io/hc/en-us/articles/360059603231-Censys-Internet-Scanning-Intro#h_01F38GFG3047EFZ4YDJDTYWJAR

TLS Fields

TLS is a service-agnostic cryptographic protocol, so the Censys schema reflects that. TLS data for any service that is using it is located at the root of the service record.

<https://search.censys.io/search/definitions?resource=hosts#TLS>

🔍 TLS

| Path | Type | Docs |
|---|--------|--|
| services.certificate | text | |
| services.jarm | object | |
| services.jarm.cipher_and_version_fingerprint | text | The first 30 byte portion of the Jarm fingerprint. |
| services.jarm.fingerprint | text | The 62 byte Jarm fingerprint of the service. |
| services.jarm.observed_at | date | The time the service was fingerprinted |
| services.jarm.tls_extensions_sha256 | text | The second 32 byte portion of the Jarm fingerprint |
| services.tls | object | |
| services.tls.certificate | object | |
| services.tls.certificate.added_at | date | When the certificate was added to the Censys dataset. |
| services.tls.certificate.ct | object | |
| services.tls.certificate.ct.entries | nested | |
| services.tls.certificate.ct.entries.key | text | |
| services.tls.certificate.ct.entries.value | object | |
| services.tls.certificate.ct.entries.value.added_to_ct_at | date | An RFC-3339-formatted timestamp indicating when the certificate was entered into the CT log. |
| services.tls.certificate.ct.entries.value.ct_to_censys_at | date | An RFC-3339-formatted timestamp indicating when the certificate was ingested from |

Note: Service names do not reflect the use of TLS (e.g., HTTPS). Search the extended service name (services.extended_service_name) to distinguish between services using or not using TLS.

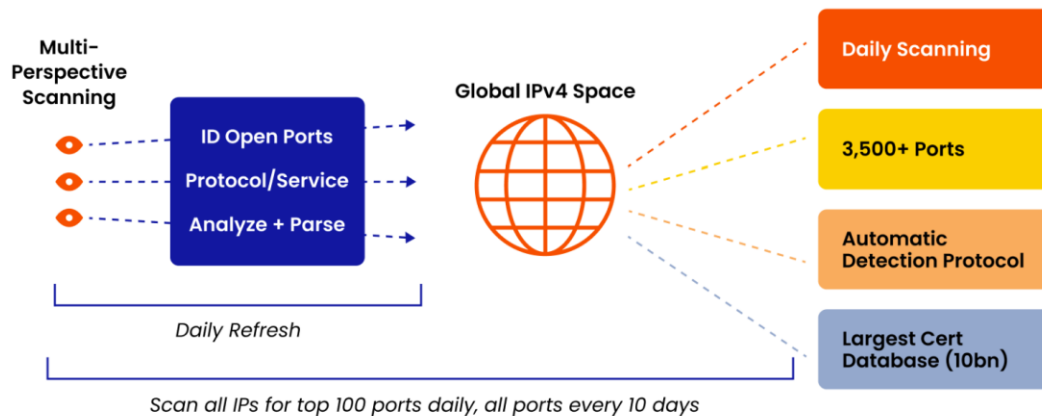
Software Fields

Within each service object, a software array shows software information in the Common Platform Enumeration (CPE) version 2.3 format. Learn more about CPE here.

</> Software

| Path | Type | Docs |
|--|---------|---|
| services.software | nested | |
| services.software.component_uniform_resource_identifiers | text | URIs of software components related to the identified software. |
| services.software.cpe | text | CPE uri format as defined here: https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7695.pdf |
| services.software.edition | text | Captures edition-related terms applied by the vendor to the product, deprecated in CPE 2.3, but kept for backwards compatibility with CPE 2.2. |
| services.software.language | text | Valid language tag as defined by [RFC5646], and should be used to define the language supported in the user interface of the product being described. |
| services.software.other | object | Other attributes describing the identified software |
| services.software.other.key | text | |
| services.software.other.value | text | |
| services.software.part | keyword | Defines the class of this software, a for application, o for operating system, h for hardware devices. |
| services.software.product | text | Identifies the most common and recognizable title or name of the product. |
| services.software.source | text | Defines the source that this software information was derived from. |
| services.software.sw_edition | text | Characterizes how the product is tailored to a particular market or class of end users. |
| services.software.target_hw | text | Characterizes the instruction set architecture (e.g., x86) on which the product being described. Bytecode-intermediate languages, such as Java bytecode for the Java Virtual Machine or Microsoft Common Intermediate Language for the Common |

How Censys Search Works



How to start with Censys search:

Censys Search Language (CSL) is a query language specifically designed for searching and retrieving information from the Censys database, which is a widely-used search engine for Internet-wide scanning and data collection. Censys collects data on hosts and websites through various scanning techniques such as port scanning, TLS certificate parsing, and DNS interrogation. CSL allows users to craft complex queries to search for specific information within this vast dataset.

Introduction:

Censys Search Language (CSL) is a powerful query language designed to extract specific information from the extensive dataset maintained by Censys, a popular search engine for Internet-wide scanning and data collection. CSL provides users with the ability to construct complex queries to search for hosts, certificates, and other relevant data based on various criteria.

Overview of CSL:

CSL enables users to construct queries using structured syntax, operators, and filters to refine search results. It supports full-text searches, field-value searches, wildcard searches, boolean logic searches, nested searches, and ranges, providing flexibility and precision in querying the Censys database.

Key Concepts and Types of searches:

- 1) **Full Text Searches:** Users can search for specific terms or phrases across all text-based fields within the dataset. Queries can be case-insensitive and support multi-word phrases enclosed in double quotes for precise matching.
➔ A query that doesn't specify a field search across all text-based fields for the word or phrase submitted.

Example:

Query for hosts with any field that contains the term "security":

The screenshot shows the Censys search interface with the query "security". The left sidebar contains "Host Filters" with categories like Labels, Autonomous System, and Location. The main area displays "Hosts" results, showing details for two IP addresses: 152.67.19.115 and 152.67.178.224. Each host entry includes its operating system (Linux), autonomous system (ORACLE-BMC-31898), location (Maharashtra, India and Telangana, India), and a list of open ports and services.

You can search for a multiple-word phrase by surrounding it in double quotes.

Example:

Query for Hosts with any field that contains the phrase “cyber security”:

The screenshot shows the Censys search interface with the query "cyber security". The left sidebar contains "Host Filters" with categories like Labels, Autonomous System, and Location. The main area displays "Hosts" results, showing details for three IP addresses: 141.148.75.149, 132.226.152.132, and 103.147.4.143. Each host entry includes its operating system (Linux), autonomous system (ORACLE-BMC-31898 and IDNIC-SIGMA-AS-ID PT. SIGMA CIPTA CARAKA), location (Virginia, United States, Arizona, United States, and Jakarta, Indonesia), and a list of open ports and services.

Note: These searches are not case sensitive. You see results with any capitalization of the letters in your search term.

- 2) **Field-Value Searches:** CSL allows searching structured fields using dot notation to specify nested keys. Users can find records based on specific values stored in those fields, such as software products, IP addresses, ports, etc.

Search structured fields for a value stored there. Fields reflect the nested structure of the host schema using dot notation to separate keys.

Find all devices that have a software product with the word "Mac" in it:

The screenshot shows the Censys search interface. The search bar at the top contains the query 'services.software.product: Mac'. The results page displays a list of hosts. On the left, there are filters for 'Host Filters' (Labels: login-page, remote-access, file-sharing, tarpit, email) and 'Autonomous System' (AMAZON-02, AMAZON-AES, DTAG Internet service provider operations, UUNET, ATT-INTERNET4). The main results section shows three hosts: 70.158.118.5 (austria.pageplanet.com), 162.155.254.131 (162-155-254-131.biz.spectrum.com), and 70.158.118.13 (australia.pageplanet.com). Each host entry includes details like 'Apple Mac Os', 'BELLSOUTH-NET-BLK (6389)', and 'North Carolina, United States'. The bottom right corner shows a notification badge with the number 81.

Important: The search above does not limit hits to the exact word specified. So a host with “Apple Mac OS X” in the server header is also returned as a result.

- If you want to search for an exact match (that is, only the word "Windows" and nothing more), replace the colon between the field name and value with an equals sign (=).

Example:

Find all devices whose software product is parsed as exactly the word "Windows:"

services.software.product=Windows

censys KS

Results [Try CensysGPT Beta](#) [Report](#) [Docs](#) [Subscriptions](#)

Host Filters

Labels:

- 1.93M remote-access
- 1.72M network-administration
- 1.31M file-sharing
- 551.15K database
- 468.03K login-page
- [More](#)

Autonomous System:

- 1.02M MICROSOFT-CORP-MSN-AS-BLOCK
- 247.42K TENCENT-NET-AP Shenzhen Tencent Computer Systems Company Limited
- 240.57K ALIBABA-CN-NET Hangzhou Alibaba Advertising Co.,Ltd.
- 187.23K KIXS-AS-KR Korea Telecom
- 134.86K AMAZON-02

Hosts

Results: 5,818,572 Time: 0.50s

92.154.54.78 (laubervilliers-659-1-35-78.w92-154.abo.wanadoo.fr)

Microsoft France Telecom - Orange (3215) Île-de-France, France

file-sharing network.device.firewall remote-access network.device

21/FTP >_ 22/SSH 80/HTTP 443/HTTP 444/HTTP

180.101.181.247

Microsoft Windows CHINANET-BACKBONE No.31,Jin-rong Street (4134) Jiangsu, China

network-administration remote-access

135/DCERPC 808/HTTP 1080/SOCKS 3389/RDP 47001/HTTP

49.235.224.240

Microsoft Windows TENCENT-NET-AP Shenzhen Tencent Computer Systems Company Limited (45090) Shanghai, China

jquery bootstrap login-page remote-access network-administration

80/HTTP 135/DCERPC 139/NETBIOS 1801/MSMQ 3389/RDP

4369/EPMD 5357/HTTP 5672/AMQP 8080/HTTP 15672/HTTP

- 3) **Wildcard Searches:** Wildcard symbols (*) and (?) enable users to search for records with unknown parts in the value. This is useful for finding variations of terms or when the exact value is not known.

censys KS

Results [Report](#) [Docs](#) [Sub](#)

Host Filters

Labels:

- 63.47M ipv6
- 11.74M remote-access
- 5.32M network-administration
- 3.34M proxy
- 3.28M file-sharing
- [More](#)

Autonomous System:

- 61.99M AMAZON-02
- 6.89M BT-UK-AS BTnet UK Regional network
- 4.58M AS-HOSTINGER
- 3.46M AMAZON-AES
- 3.03M GOOGLE-CLOUD-PLATFORM
- [More](#)

Location:

- 83.83M United States
- 47.37M United Kingdom

Hosts

Results: 131,526,242 Time: 1.11s

35.88.173.114

Lancom Systems AMAZON-02 (16509) Oregon, United States

printer remote-access tarpit truncated

21/HTTP >_ 22/SSH 113/HTTP 443/HTTP 2052/HTTP

>_ 2222/SSH 3111/HTTP 4100/HTTP 4444/HTTP 4840/HTTP

5060/HTTP 5222/HTTP 5556/HTTP 5713/HTTP 6000/UNKNOWN

6007/HTTP 8048/HTTP 8089/HTTP 8143/HTTP 8174/HTTP

8446/HTTP 8859/HTTP 9047/HTTP 9055/HTTP 9070/HTTP

As well as 44 more

154.30.236.109

AS-COLOCROSSING (36352) California, United States

proxy

21230/HTTP 21231/HTTP 21234/HTTP 21235/HTTP 21236/HTTP

21237/HTTP 21239/HTTP 21240/HTTP 21241/HTTP 21242/HTTP

21243/HTTP 21244/HTTP 21246/HTTP 21248/HTTP 21249/HTTP

21250/HTTP 21252/HTTP 21253/HTTP 21254/HTTP 21255/HTTP

21259/HTTP 21260/HTTP 21263/HTTP 21265/HTTP 21266/HTTP

Note: Use the asterisk symbol (*) to substitute zero or more unknown characters.

Use the question mark (?) to substitute for exactly 1 unknown character.

- 4) **Boolean Logic Searches:** CSL supports logical operators (AND, OR, NOT) and parentheses to combine multiple search criteria effectively, allowing for complex queries to be constructed.

- **OR**

Use or to provide multiple options that a record can match to be considered a hit.

Example:

Return all hosts located either in the Canada or France:

location.country: Canada or location.country: "France"

The image displays two screenshots of the Censys search interface, demonstrating the results of a Boolean search query: `location.country: Canada or location.country: "France"`.

Top Screenshot (France):

- Search Query:** `location.country: UK or location.country: "France"`
- Results:** 15,450,540
- Hosts:**
 - 51.158.61.218** (3144bdd9-2ad9-40d4-9e32-99f9a51d451b.fr-par-2.baremetal.scw.cloud)
 - OS: Linux, Provider: Online SAS (12876), Location: Île-de-France, France
 - Tags: bootstrap, jquery, jquery-migrate, jquery-ui, lot, database, remote-access, email, login-page
 - Services: 22/SSH, 443/HTTP, 4559/SMTP, 25/SMTP, 993/IMAP, 5038/UNKNOWN, 80/HTTP, 995/POP3, 5353/MDNS, 110/POP3, 3306/MYSQL, 143/IMAP, 4190/UNKNOWN, 8088/HTTP, 8089/HTTP
 - 176.31.249.156** (ns387950.ip-176-31-249.eu)
 - OS: Linux, Provider: OVH (16276), Location: Hauts-de-France, France
 - Tags: authentication, web.control-panel.hosting, remote-access, email, file-sharing
 - Services: 21/FTP, 106/POPPASSD, 993/IMAP, 8443/HTTP, 22/SSH, 110/POP3, 995/POP3, 8880/HTTP, 25/SMTP, 143/IMAP, 4190/PIGEONHOLE, 7080/HTTP, 53/DNS, 443/HTTP, 465/SMTP, 7081/HTTP
 - 213.32.64.171** (171.ip-213-32-64.eu)
 - OS: Linux, Provider: OVH (16276), Location: Hauts-de-France, France
 - Tags: email, file-sharing, remote-access, web.control-panel.hosting, prototype, requires

Bottom Screenshot (Canada):

- Search Query:** `location.country: Canada or location.country: France`
- Results:** 24,804,531
- Hosts:**
 - 45.60.206.61**
 - Provider: INCAPSULA (19551), Location: Ontario, Canada
 - Tags: truncated
 - Services: 53/DNS, 21/HTTP, 631/IPP, 4510/UNKNOWN
 - 45.60.100.176**
 - Provider: INCAPSULA (19551), Location: Ontario, Canada
 - Tags: truncated
 - Services: 53/DNS, 21/HTTP, 631/IPP, 14764/UNKNOWN
 - 174.115.181.85**
 - Provider: ROGERS-COMMUNICATIONS (812), Location: Ontario, Canada
 - Tags: 7547/CWMP
 - 70.38.47.231**
 - OS: Linux, Provider: IWEB-AS (32613), Location: Quebec, Canada
 - Tags: remote-access, web.control-panel.hosting, database, file-sharing, open-dir, login-page, email

You can also use a set to shorten what can be a long or statement.

Example:

Return all hosts whose country is among this set of countries in the Europe:

location.country: {France, Germany, Italy, Switzerland, Sweden, Netherland}

The screenshot shows the Censys search interface. At the top, the search bar contains the query `location.country: {France, Germany, Italy, Switzerland, Sweden, Netherland}`. The search results are displayed in a table with columns for Hosts, Labels, Autonomous System, and Location. The first result is for IP `87.245.101.205` (kerio.wolfkeller.ch), which is associated with Redhat Enterprise Linux, SASAG (35518), and Schaffhausen, Switzerland. The second result is for IP `77.56.67.143` (77-56-67-143.dclient.hispeed.ch), associated with LIBERTYGLOBAL Liberty Global formerly UPC Broadband Holding, aka AORTA (6830), and Zurich, Switzerland. The third result is for IP `46.14.148.202` (202.148.14.46.static.wline.lns.sme.cust.swisscom.ch), associated with SWISSCOM Swisscom Switzerland Ltd (3303), and Zurich, Switzerland. The fourth result is for IP `45.223.220.251`, associated with INCAPSULA (19551), and Zurich, Switzerland.

- AND

Use and to make a search more specific by providing multiple criteria that must match for a host to be considered a hit.

Example:

Return hosts with port 80 open (with any service type) and an HTTP service (on any port):

services.port: 80 and services.service_name: HTTP

censys Hosts services.port: 80 and services.service_name: HTTP Search KS

Hosts List:

- 151.68M ipv6
- 16.25M remote-access
- 8.21M file-sharing
- 6.38M login-page
- 5.83M jquery
- More

Autonomous System:

- 150.13M AMAZON-02
- 5.56M AKAMAI-AS
- 4.14M AS-HOSTINGER
- 1.82M AMAZON-AES
- 1.65M MICROSOFT-CORP-MSN-AS-BLOCK
- More

Location:

- 87.98M India
- 39.70M United States
- 28.25M United Kingdom
- 5.69M Germany
- 5.51M Brazil
- More

Service Filters

Service Names:

Host Details:

- 103.248.60.254** (Microsoft Windows, WEBWERKS-AS-IN Web Werks India Pvt. Ltd. (133296), Maharashtra, India)
 - database, prototype, requirejs, file-sharing, email
 - 21/FTP, 25/SMTP, 53/DNS, 80/HTTP, 110/POP3
 - 135/DCERPC, 139/NETBIOS, 143/IMAP, 443/UNKNOWN, 445/SMB
 - 465/SMTP, 993/IMAP, 995/POP3, 1433/MSSQL, 3306/MYSQL
 - 5443/HTTP, 5985/HTTP, 8306/MYSQL, 8443/HTTP, 8880/HTTP
 - 18018/HTTP, 47001/HTTP
- 23.198.186.81** (a23-198-186-81.deploy.static.akamaitechnologies.com)
 - Akamai, AKAMAI-AS (16625), Buenos Aires F.D., Argentina
 - 80/HTTP, 443/HTTP
- 172.105.56.60** (172-105-56-60.ip.linodeusercontent.com)
 - Linux, AKAMAI-LINODE-AP Akamai Connected Cloud (63949), Maharashtra, India
 - file-sharing, login-page, web.control-panel.hosting, email, bootstrap, google-analytics, jquery, remote-access
 - 21/FTP, 25/SMTP, 53/DNS, 80/HTTP, 110/POP3
 - 143/IMAP, 443/HTTP, 465/SMTP, 587/SMTP, 993/IMAP
 - 2077/HTTP, 2078/HTTP, 2082/HTTP, 2083/HTTP, 2086/HTTP
 - 2087/HTTP, 2095/HTTP, 2096/HTTP, >... 30000/SSH
- 2606:4700:3032:0:0:ac43:c8f0**
 - CLOUDFLARENET (13335), California, United States
 - ipv6

- NOT

Use not to exclude hosts with certain characteristics.

Example:

Find hosts running HTTP on a non-standard port (i.e., neither 22 nor 2222):

services: (service_name: HTTP and not port: {22, 2222})

censys Hosts services: (service_name: HTTP and not port: {22, 2222}) Search KS

Results Try CensysGPT Beta Report Docs Subscriptions

Host Filters

Labels:

- 197.53M ipv6
- 26.67M remote-access
- 13.40M login-page
- 12.67M jquery
- 10.45M network.device
- More

Autonomous System:

- 191.38M AMAZON-02
- 10.75M AS-HOSTINGER
- 7.55M KIXS-AS-KR Korea Telecom
- 6.84M BT-UK-AS BTnet UK Regional network
- 5.67M ASN-IBSNZ
- More

Location:

- 109.95M India
- 72.76M United States

Hosts

Results: 371,975,578 Time: 4.07s

- 198.143.37.18** (198.143.37.18.ip.incapdns.net)
 - INCAPSULA (19551), Virginia, United States
 - truncated
 - 53/DNS, 21/HTTP, 631/IPP, 9127/UNKNOWN
- 185.123.141.73**
 - Linux, M247 (9009), București, Romania
 - camera
 - 80/HTTP, 443/HTTP, 554/RTSP, 2905/UNKNOWN, 49152/HTTP
- 198.143.52.4** (198.143.52.4.ip.incapdns.net)
 - INCAPSULA (19551), Victoria, Australia
 - truncated
 - 53/DNS, 21/HTTP, 631/IPP, 13294/UNKNOWN
- 107.154.37.242** (107.154.37.242.ip.incapdns.net)
 - INCAPSULA (19551), California, United States

- 5) **Nested Searches:** Users can apply search criteria to specific objects within a list of similar objects, such as services on a host. This enables more granular filtering and targeting of search results.

Use nested query syntax to apply multiple search criteria to a single object within a list of like objects, instead of to the entity as a whole.

Example:

Return hosts with an HTTP service returning a specific Etag header value:

services.http.response.headers: (key: 'Etag' and value.headers: '"6001043d.16d"')

The screenshot shows the Censys search interface. The search bar contains the query: `services.http.response.headers: (key: 'Etag' and value.headers: '"6001043d.16d"')`. The results are displayed in a table with columns for Hosts, Services, and Headers. The table lists three hosts: 113.147.227.31, 101.143.218.148, and 118.6.184.236. Each host is associated with a specific organization and location. The table also shows various services (e.g., 161/SNMP, 50001/HTTP) and their corresponding headers (e.g., 60001/HTTP, 60002/HTTP).

- 6) **Ranges:** Ranges allow users to define a spectrum for numerical values like dates, version numbers, and IP addresses, facilitating searches within specific intervals.

➔ Search for online hosts in a non-standard IP address range, including the first and last values given:

ip: [119.167.243.56 to 119.167.243.201]

The screenshot shows the Censys search interface. The search bar contains the query: `ip: [119.167.243.56 to 119.167.243.201]`. The results are displayed in a table with columns for Hosts, Services, and Headers. The table lists three hosts: 119.167.243.65, 119.167.243.56, and 119.167.243.64. Each host is associated with a specific organization and location. The table also shows various services (e.g., 80/HTTP, 443/HTTP) and their corresponding headers (e.g., 443/HTTP, 444/HTTP).

Note: Censys also supports CIDR notation of IP ranges, which shows a IPv4 or IPv6 address with a slash followed by a decimal to indicate how many bits in binary identifiers (from left to right) are fixed and do not change.

Example:

Search for online hosts whose IPv4 address fall between 35.180.0.0 and 35.180.255.255:

ip: 35.180.0.0/16

The screenshot shows the Censys search interface. The search bar contains the query 'ip: 35.180.0.0/16'. The results page displays a list of hosts under the 'Hosts' section. The first host is 35.180.67.109, which is part of the AMAZON-02 (16509) cloud provider and is located in Île-de-France, France. It lists various services such as 53/DNS, 445/SMB, 523/DB2, 10001/UBIQUITI, 53413/NETIS, 143/IMAP, 5353/MDNS, 5351/NATPMP, 3702/WS_DISCOVERY, 1883/MQTT, 1604/CITRIX, 137/NETBIOS, 69/TFTP, 2/HTTP, 5683/COAP, 5432/POSTGRES, 5093/SENTINEL, 17185/WDBRPC, 25/SMTP, 3283/ARD, 19/CHARGEN, 17/QOTD, 22/SSH, 177/XDMCP, and 3389/RDP. The second host is 35.180.173.47, also part of AMAZON-02 (16509) and located in Île-de-France, France. It lists services like 69/TFTP, 26257/POSTGRES, 5093/SENTINEL, 25/SMTP, 1080/SOCKS, 53/DNS, 19/CHARGEN, 523/DB2, 8883/MQTT, 6379/REDIS, 143/IMAP, 53413/NETIS, 27017/MONGODB, 5683/COAP, 520/RIPV1, 22/SSH, 5353/MDNS, 37215/UPNP, 3283/ARD, 631/IPP, 1604/CITRIX, 137/NETBIOS, 2/HTTP, 3389/RDP, and 23/UNKNOWN. The left sidebar shows 'Host Filters' with labels like remote-access, voip, truncated, jquery, bootstrap, and autonomous system: 17.64K AMAZON-02. It also shows 'Service Filters' with names like HTTP, SSH, SIP, XMPP, and SOCKS.

- 7) **Quotes:** Double quotes (") are used to search for exact phrases, while backticks (`) are used to escape reserved characters within certain field values, such as CPE-formatted software strings.

The screenshot shows the Censys search interface with the query 'services.http.response.html.title: "security"'. The results page displays a list of hosts under the 'Hosts' section. The first host is 156.250.223.181, which is part of the POWERLINE-AS-AP POWER LINE DATACENTER (132839) and is located in Hong Kong. It lists services like 80/HTTP, 443/HTTP, 888/HTTP, 25441/SSH, and 59361/HTTP. The second host is 3.24.229.14 (ec2-3-24-229-14.ap-southeast-2.compute.amazonaws.com), which is part of the AMAZON-02 (16509) cloud provider and is located in New South Wales, Australia. It lists services like 22/SSH, 80/HTTP, 111/PORTMAP, 5003/TELNET, 80/HTTP, bootstrap, jquery, owl-carousel, slick, 443/HTTP, 1150/HTTP, 7757/HTTP, and 12618/SSH. The third host is 168.76.15.65, which is part of the ASLINE-AS-AP ASLINE LIMITED (137951) and is located in Hong Kong. It lists services like 80/HTTP, 18598/HTTP, 443/HTTP, 1150/HTTP, 7757/HTTP, and 12618/SSH. The fourth host is 63.84.48.195 (autodiscover.cpoliceservice.com). The left sidebar shows 'Host Filters' with labels like network.device, remote-access, login-page, file-sharing, network-administration, and autonomous system: 6,785 ATT-INTERNET4, 6,684 ZSCALER-SJC1, 5,640 UUNET, 5,517 CMCS, 4,963 ZSCALER-EMEA. It also shows 'Location' with counts for United States (95.11K), United Kingdom (8,693), India (8,242), and Japan (8,165).

Black ticks(`)

Backticks escape all reserved characters occurring therein. For example, CPE-formatted software strings use many reserved characters. Instead of escaping each one, wrap the whole string in backticks.

Search for hosts running Microsoft IIS version 10.0:

```
services.software.uniform_resource_identifier:`cpe:2.3:o:microsoft:windows_server_2012:*:*:*:*:*:*:*`
```

The screenshot shows the Censys search results for the query: `services.software.uniform_resource_identifier:`cpe:2.3:o:microsoft:windows_server_2012:*:*:*:*:*:*:*``. The results are displayed in a table-like format with three main host entries:

- Host 1:** 101.99.36.93 (static.cmcti.vn). Organization: CMCTELECOM-AS-VN CMC Telecom Infrastructure Company (45903). Location: Ho Chi Minh, Vietnam. Services: remote-access, login-page, default-landing-page, network.device.vpn, network.device. Ports: 443/HTTP, 1194/OPENVPN, 1701/L2TP, 8069/HTTP, 8079/HTTP, 59101/HTTP, 59105/HTTP.
- Host 2:** 223.108.130.2. Organization: CMNET-JIANGSU-AP China Mobile communications corporation (56046). Location: Jiangsu, China. Services: login-page, jquery, iot. Ports: 83/HTTP, 1883/MQTT, 2026/HTTP, 2028/HTTP, 4343/HTTP, 4430/HTTP, 4433/HTTP, 8088/HTTP, 9443/HTTP, 11000/HTTP, 11198/HTTP, 11199/HTTP.
- Host 3:** 154.92.23.165. Organization: YISU CLOUD LTD (138152). Location: Virginia, United States. Services: login-page, file-sharing, network-administration, jquery, jquery-ui, database, default-landing-page, remote-access. Ports: 21/FTP, 80/HTTP, 135/DCERPC, 137/NETBIOS, 139/NETBIOS, 2206/HTTPS, 2280/HTTPS, 5085/HTTPS, 5588/HTTPS.

Introduction to Virtual Hosts

Hosts are identified by IP address. Virtual hosts are identified by a name + IP address.

Collecting Virtual Host Data with Name-based Scans

Virtual hosts contain services that responded to a name-based Censys scan. The name of the virtual host is the name used to scan its services.

This name is included in the scan in 1 of 2 ways:

- ❖ In the server name indicator (SNI) field during a TLS handshake.
- ❖ In the Host header field of an HTTP request.

Virtual hosts do not have the same top-level information that hosts do (e.g., geographic location, routing data, DNS data, etc.), with the exception of operating system and labels.

Services on Virtual Hosts

Virtual host records include an array of services (that responded in scan to the virtual host's name) with the same fields as those seen on hosts.

The names of services on virtual hosts are limited compared to those seen on hosts because many protocols do not support name-based differentiation.

Values for `services.service_name` that can appear on a virtual host:

- ANYCONNECT
- HTTP
- ELASTICSEARCH
- KUBERNETES
- PROMETHEUS
- UNKNOWN

The Overlap of Hosts and Virtual Hosts

Service data is not necessarily different between an (unnamed) host and a virtual host. A service observed on a particular port may provide the same content when a name is specified and when it is not.

Some services are only seen when a name is specified, so a virtual host can have whole services that are not present on the underlying (unnamed) host.

The Effect of Virtual Hosts on Search Results

The number of vhosts in the Censys dataset is more than double the number of hosts. Including virtual hosts in your search results can dramatically increase the number of hits.

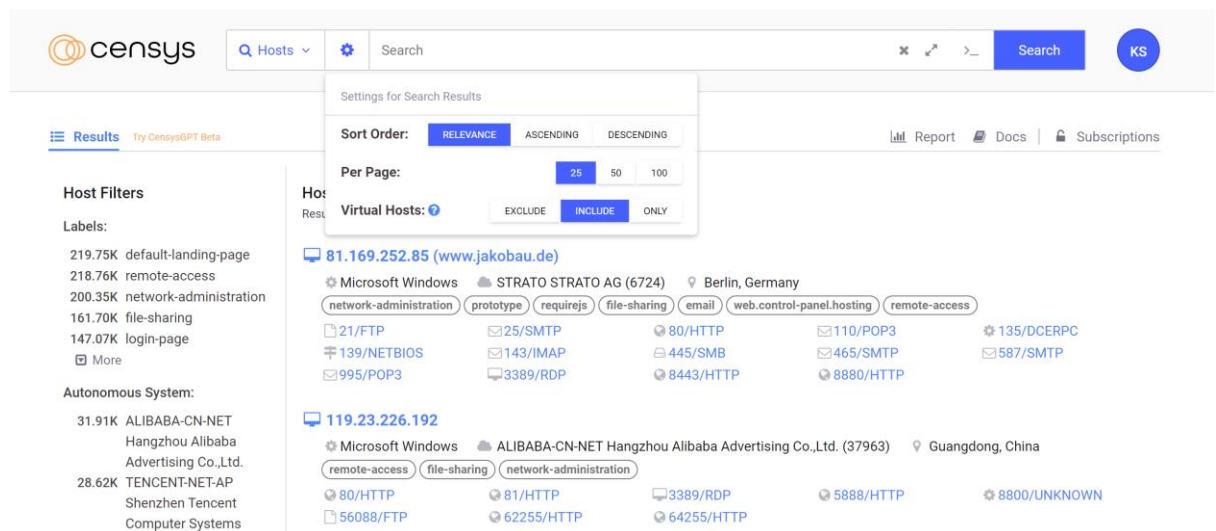
Due to the volume of virtual host services in the Censys dataset and out of respect for the integrity of hosts serving a large number of virtual hosts, name-based services are refreshed at a rate of every 30 days.

Refreshing Virtual Host Data

DNS names known to Censys are re-resolved every 30 days and successful responses result in name-based scans. Virtual hosts age out of the host index after 45 days.

How to Search Virtual Hosts in the Web UI

- 1) Click the Search Settings icon (gear) in the search bar.
- 2) To specify searching Virtual hosts, do one of the following:
 - To include virtual hosts in your search results, click Include.
 - To only search virtual hosts, click Only.
- 3) Write a query in the Censys Search Language that asks a question about virtual hosts.
- 4) Click Search.



How to Search Virtual Hosts in the API

Use the new `virtual_host` query parameter to include, exclude, or only search virtual hosts.

If not provided, the results list defaults to (unnamed) hosts only.

Example Search for Host and Vhosts Running an HTTP Service:

```
GET https://search.censys.io/api/v2/hosts/search?q=service.service_name%3A%20HTTP&per_page=1&virtual_hosts=include
```

Note: Normal hosts typically have one-to-one mappings between IP addresses and domain names. Each IP address is associated with a single domain or website.

Virtual hosts are configured to serve multiple domain names or websites from the same IP address. The web server distinguishes between different domains based on the Host header sent by the client in the HTTP request.

Introduction to Certificates

Certificates are an important part of Internet traffic encryption because they can verify the identities of the services that are communicating to each other. Censys collects certificates in a repository for searching and viewing.

The Censys certificates data set is the most exhaustive collection of X.509 documents in existence (~10B and growing daily).

Each Censys cert record contains:

- Data parsed from the certificate using ZCrypto: an open-source, Go-based, cryptographic library.
- Trust information from major root stores such as Apple, Google Chrome, Microsoft, and Mozilla NSS.
- Submission information from Certificate Transparency (CT) logs. Learn more about certificate transparency and its effect on the Censys Certificates repository.
- Lint results describing non-conformance to the X.509 standard using the ZLint library.
- Data about Censys collection and observation during scan.

Certificate Collection at Censys

Certificates are collected using 2 methods:

- 1) Syncing with a number of CT logs.
- 2) Observing a certificate presented as part of a TLS handshake during a Censys scan of the public Internet (over any protocol).

Parsed Certificate Data

The contents of a certificate are immutable and cannot be changed after the certificate is generated.

Censys parses the contents of each certificate and provides them as searchable fields in an object called `parsed`.

Parsed Fields from a Certificate (Not an exhaustive list. See all fields here)

- Issuer DN: Information about the certificate authority that issued the certificate.
- Subject DN: Information about the entity that was issued the certificate.
- Extensions: Additional fields that extend the X.509 spec.
- Validity Period: The dates from which and to which the certificate can be used.
- Serial Number: The issuer-specific identifier of the certificate.
- Public Key: The public key of the key pair that is associated with the certificate.
- Signature Algorithm: The algorithm used to sign the certificate.
- Signature Value: Bit string containing the digital signature.

Other Certificate Data

Other data about the certificate and the collection process also appear in a Censys certificate record, such as:

- Trust and validation (validation) : Information about the status of the certificate's trust by modern web browsers.
- Certificate transparency (ct) : Information about submissions to CT logs.
- Zlint (zlint): Whether the certificate's attributes triggered any lints for non-conformance to the X.509 standard.
- Seen in Scan (ever_seen_in_scan) : Whether the certificate has ever been seen during a Censys scan of the Internet. This is a one-way boolean. If true, it remains that way.

Certificate Trust and Validation

Trust chains are an important part of certificate usage. For a certificate to be trusted, the certificate must chain up, through a series of signing certificates, to a root certificate that is present in a major root trust store.

Censys indexes certificate trust information for each root store in a record called validation.

➔ Certificate Validation Fields For Each Root Store

- Valid (is_valid): A boolean value for whether the certificate is trusted by the browser using the root store.
- Ever Valid (ever_valid) : A boolean value for whether an expired certificate was trusted by the browser before it expired.
- Parents (parents): A list of the fingerprints of the intermediary and root certificates in the chain.
- Chain (chains): A representation of the chain(s) of signing certificates up to the root.
- Had Trusted Path (had_trusted_path) : A boolean value for whether the chain was trusted by the browser.
- In Revocation Set (in_revocation_set) : Whether the certificate is included in the browser's list of certs whose trust has been revoked before their expiration.

Censys regularly validates unexpired certificates. Values of validation fields and related labels are accurate as of the validated_at timestamp in the certificate record.

ZLint

Censys uses the open-source ZLint tool to lint each certificate in its collection for conformance to X.509 standards.

Lack of conformity to a specification can result in the following types of triggered lints:

- Notice
- Warning
- Error
- Fatal

Advanced Search Methods and Queries

Exploring Threat Activity & Other Interesting Artifacts

1) Open directories:

`services.http.response.html_title: "Index of/"`

2) Cobalt Strike Beacons:

`services.cobalt_strike: *`

3) Compromised MikroTik Routers:

`services.service_name: MIKROTIK_BW and "HACKED"`

4) Services on port 53 that are not DNS:

`services: (port: 53 and not service_name: DNS) and services.truncated: false`

5) Network devices with exposed login pages:

`services: (labels: {network.device, login-page})`

Command and Control infrastructure:

Deimos C2: this is the most widely used tool by pentesters

`services: (services.port: 8443 and (http.response.html_title="Deimos C2" or
tls.certificates.leaf_data.subject.organization="Acme Co"))`

Posh C2:

`services.tls.certificate.parsed.subject_dn: "C=US, ST=Minnesota, L=Minnetonka, O=Pajfds,
OU=Jethpro, CN=P18055077"`

Incident Response: Queries for a Zero-Day

Censys Search is a valuable tool when responding to a zero-day vulnerability disclosure. When a zero-day hits, the Censys Research team deploys rapid response articles that explain the scope and impact of the attack. We also include the queries that you can run to determine if you are affected.

For example, during this year's MOVEit CVE, you could run this query on hosts to identify potentially vulnerable assets:

`services.http.response.favicons.md5_hash=af8bf513860e22425eff056332282560`

Additional examples of queries to find services affected by zero-days include:

1) CVE-2023-20198 Cisco IOS-XE

labels='cisco-xe-webui'

2) CVE-2023-44487 HTTP/WHO?

services.http.supports_http2: true

3) CVE-2023-30799 MikroTik RouterOS

services.http.response.html_title: "RouterOS router configuration page"

Meta/Facebook Pixel Trackers

Customers have also recently used Censys Search to determine the presence of Meta Pixel Javascript code, which tracks website user activity through cookies and sends information back to Meta, which can pose a problem for sensitive data.

You can use the following query to determine the presence of Meta Pixel code on their websites:

services.http.response.body:"fbq('track', 'PageView');"

Unmasking Deception: Honeypot and Red Herring in Network Security

Censys main mission is to craft the ultimate blueprint of the web, map all the strange anomalies, and unearth where the wild things roam. It scans the internet indiscriminately and does an excellent job, too. And when you look at this data all day, you tend to become accustomed to the strange little quirks (like commercial honeypots) you often encounter and become desensitized to the extremely odd things.

One reality that quickly emerges from this data is realizing the internet's abundance of deception. There are no laws or regulatory mechanisms to compel internet-connected hosts to disclose their true identity or purpose: For example, you can generate an SSL certificate for google.com, and virtually no one is in your way to prevent you from deploying it. However, most hosts are generally unlikely to deceive us about their true nature, as crafting and maintaining these falsehoods requires some effort.

- In Censys, a honeypot is a simulated computer or server deployed on the internet to attract cyber attackers. It's like a digital trap designed to gather information about hackers' activities.

- Censys uses fake servers called honeypots to trick hackers into attacking them. These honeypots pretend to be vulnerable systems, enticing attackers to try to break in. When attackers interact with the honeypots, Censys records their actions to learn more about their tactics.

- Censys uses honeypots to study cyber threats and improve security measures. By analyzing hackers' behavior in these controlled environments, Censys can better understand the risks and vulnerabilities present on the internet.

- Censys deploys honeypots across the internet, strategically placing them in locations where attackers are likely to target. These honeypots mimic various types of servers and services to attract different types of attacks.

- While honeypots in Censys can be effective, they also face challenges. Some honeypots may be easily detected by attackers, reducing their effectiveness. Additionally, distinguishing between legitimate and malicious activity within the honeypots can be difficult.

For example, the specialized honeypot software GasPot attempts to emulate a legitimate Automated Tank Gauging service (ATG) (used for monitoring fuel levels) but is easily unmasked with little scrutiny.

```
mark@bleep:~/pipeline_current$ echo AUkyMDEwMAo= | base64 -d | nc 18.141.13.188 10001
I20100
OCT 5, 2023 10:57 AM
861702 V
IN-TANK INVENTORY
TANK PRODUCT VOLUME TC VOLUME ULLAGE HEIGHT WATER TEMP
1 PREMIUM 2028 2007 5706 29.01 0.00 74.47
2 UNLEADED 2 6467 6395 3053 58.58 0.83 75.78
3 UNLEADED 7385 7309 2135 65.87 0.00 74.55
4 DIESEL 3174 3149 6346 34.35 0.00 77.25
5 KEROSENE 667 662 335 30.32 0.82 73.62
```

(An actual ATG service)

```
mark@bleep:~/pipeline_current$ echo AUkyMDEwMAo= | base64 -d | nc 18.141.13.188 10001
I20100
10/05/2023 15:00
AVIA
IN-TANK INVENTORY
TANK PRODUCT VOLUME TC VOLUME ULLAGE HEIGHT WATER TEMP
1 SUPER 8430 8600 3792 55.90 5.77 50.01
2 UNLEAD 4971 4993 4324 41.25 6.96 55.57
3 DIESEL 5184 5289 8588 29.11 2.74 58.97
4 ADBLUE 4613 4804 8588 54.18 5.92 59.63
```

(A GasPot (Fake) ATG service)

Three indicators differentiate GasPot and an actual ATG device:

- 1) GasPot has a limited number of diagnostic codes that it will accept, and for any code it does not understand, it will return the error code "9999FF1B".
- 2) GasPot formats the timestamps in the payloads differently than real ATG devices. For example, GasPot formats them as "MM/DD/YYYY HH:MM", whereas an actual ATG device formats its timestamps like this: "Nov 8, 2022 15:45"
- 3) Real ATG devices use CRLF ("\r\n"), while GasPot primarily uses newlines ("\n\n") due to the code in the following screenshot


```
def I20100():
    I20100_1 = ''
    I20100
    '''
    I20100_2 = ''

    ''' + station + '''

IN-TANK INVENTORY

TANK PRODUCT          VOLUME TC VOLUME  ULLAGE  HEIGHT  WATER  TEMP
1  ''' + PRODUCT1 + ''' + str(Voll) + '''      ''' + str(volltc) + '''      ''' + ullage1 + '''
ol + '''      ''' + templ + '''
2  ''' + PRODUCT2 + ''' + str(Vol2) + '''      ''' + str(vol2tc) + '''      ''' + ullage2 + '''
```

The GasPot code that generates newlines instead of CRLF

With that known, it's reasonably easy to use Censys to find hosts running this GasPot honeypot server simply by searching for ATG services not using newlines instead of CRLF: `services: (service_name=ATG and banner="*\n\n\n\n")`

And when you search for these GasPot services, you will notice that the majority of the results have hosts with all sorts of “interesting” and uncommon features and classifications of services that are often not found running together in the real world.

80.156.42.78
 DTAG Internet service provider operations (3320) Bavaria, Germany
 (truncated)
 1 Matched Service
 10001/ATG
 20 Other Services
 143/IMAP 22/SSH 445/SMB 2404/IEC60870_5_104 1723/PPTP
 21/FTP 110/POP3 1883/MQTT 1433/MSSQL 27017/MONGODB
 6379/REDIS 80/HTTP 25/SMTP 5432/POSTGRES 11112/DICOM
 3306/MYSQL 5900/VNC 9200/ELASTICSEARCH 631/IPP 23/UNKNOWN

13.238.143.227 (ec2-13-238-143-227.ap-southeast-2.compute.amazonaws.com)
 AMAZON-02 (16509) New South Wales, Australia
 (truncated)
 1 Matched Service
 10001/ATG
 19 Other Services
 123/NTP 143/IMAP 22/SSH 445/SMB 2404/IEC60870_5_104
 1723/PPTP 21/FTP 3306/MYSQL 110/POP3 5900/VNC
 1433/MSSQL 27017/MONGODB 2/HTTP 25/SMTP 5432/POSTGRES
 5060/SIP 1883/MQTT 631/IPP 23/UNKNOWN

52.65.170.103 (ec2-52-65-170-103.ap-southeast-2.compute.amazonaws.com)
 AMAZON-02 (16509) New South Wales, Australia
 (truncated)
 1 Matched Service
 10001/ATG
 15 Other Services
 143/IMAP 445/SMB 2404/IEC60870_5_104 1723/PPTP 21/FTP
 3306/MYSQL 110/POP3 5900/VNC 1433/MSSQL 27017/MONGODB
 2323/UNKNOWN 80/HTTP 25/SMTP 5060/SIP 1883/MQTT

In the GasPot result screenshot above, many hosts have four or five different database technologies that are functionally identical (MSSQL, MySQL, Postgres, etc.). We also see services commonly associated with everyday web applications running alongside IoT and SCADA services. To top it all


off, many of these ATG servers live in AWS, which, to my knowledge, doesn't have direct access to physical tanks of gasoline. It's not the best representation of reality.

What are red herrings?

Network scanners like Censys will record information from a service exactly as presented by the host, and on top of the raw data, we will augment the host details with information about the running services and software using labels and CPEs. The logic behind finding and applying these software and service labels is, for the most part, a simple process involving regular expressions and pattern matching using both internal and open-source data. And for most hosts on the internet, this works perfectly fine.

So when Censys were made aware of a new set of hosts that people were talking about on social media that attempted to not only lie about who they were but seemingly try to overload network scanners with false positives, they weren't surprised as they've witnessed similar things before.

On September 20, 2023, Censys started observing around 50 hosts with a unique and chaotic characteristic: in the HTTP response, these hosts included a 37,213-byte Server header (customarily used to identify the running server) with hundreds of different software names.

| 7004/HTTP TCP | | View Definition |
|----------------------------|--|---|
| Attribute | Value | |
| services.banner | HTTP/1.1 200 OK\r\nDate: <REDACTED>\r\nServer: 979634648,-1492966240,Downloaded,-1642532491,strcorpid,ADMINCONSOLESESSION=,CardNo,PTG,Location:/ShUWsRNbOn7z07.9v6CC1_q6MF8MD0q-..HxgkLdn8G2Ee.8kjXLq1zL_ZknxpCjinteract.sh/,-1267819858,page_aajoda-testimoni als,application/java,x-cmd-response,azkaban.browser.session.id,ac_pass wd=,Server: Lotus-Domino,SonicWALL SSLVPN Web Server,905744673,{(r esult)},Sign In,Server: proscan,-219752612,application/download,Set-Cook ie: pbootsystem=-,2145085239,-41369781,Set-Cookie: rememberMe=dele teMe,Server: GateOne,1251810433,-297069493,/wp-includes/,reporter_ad min,-1148190371,708578229,Location:/fr9U36SZPTHY2m64YdIY19a6raY 5BqXI9ockcT6lyRpn234C.JHSS.i832AUmZ8qG.Minteract.sh/,251106693, app/template/Index.vm?login=true,ugm,-confluence,-RStudio,X-Cocoon-V ersion:30455281067826 53260371. 026 0793183478245 924402936 .66 50830339418985824738 7721 ,ERROR URL,Server: compaqhttpserver,sel ea_httpd,datauser,WP_SESSIONID=,text/zyxel,Server: easy file sharing web se rver,Content-Type: text/xml,41e9c43dc5e994ca7a40f4f92b50d01d,78653 3217,x-oracle-dms-ecid,{(randstr_1)},fileDownload=true,-1507567067,BEGI N RSA PRIVATE KEY,59a0c7b6e4848ccdadabcea0636efda02b,MTc5Njc2NT UwNg==,text/tab-separated-values,Server: kfwebserver,Sign In - Airflow,Se |  |

Over the next few weeks, they saw the number of hosts with this data increase dramatically, growing from three to six thousand hosts daily. By September 30th, they saw over 27,252 unique hosts presenting this huge and obnoxious server header.

More interesting is where these hosts were located (geographically and AS-wise). At the time of writing, all hosts exist in the autonomous system AMAZON-02, one of the largest AWS networks. But, at the start of this event, two other ASs were seen with these hosts: AMAZON-AES (AS14618) and BJ-GUANGHUAN-AP (AS55960), both of which stopped these specific services nine days later, on September 29th.

Beijing Guanhuan Xinwang Digital is the Internet Service Provider that Amazon partners with to legally operate AWS in China, meaning the IP blocks this AS announces are owned and operated by Amazon, but the data centers and transit are controlled by Xinwang Digital.

Given that you must have a valid Chinese business license to open an account in AWS China, the mere existence of these hosts points to a China-based operation. Interestingly, they attempted to mask this by moving everything to AMAZON-02, located primarily in the United States, late last month. With the sheer scale of this incident, it points to a very well-planned experiment with foreign, corporate (possibly state-based or even educational) backing. Since deploying services into AWS provides anonymity, and every one of these servers seems to be structurally similar without any other indication of who owns them, it is hard to tell precisely who is behind this.

Applications of Censys:

1. Industrial Control Systems (ICS) and SCADA (Supervisory control and data acquisition) Systems:

- SCADA systems are used to control industrial equipment like motors and sensors.
- Censys helps identify and categorize publicly available Modbus devices used in SCADA systems.
- Modbus devices lack security measures, posing risks to industrial infrastructure.
- Censys found Modbus devices in 117 countries, with the majority in the United States.

2. Heartbleed, Poodle, and SSLv3 Vulnerabilities:

- Heartbleed and Poodle were major vulnerabilities in 2014.
- Censys helps identify vulnerable hosts and track SSLv3 support.
- Despite disclosures, some hosts remain vulnerable to Heartbleed and support SSLv3.

3. Institutional Attack Surface Management:

- Organizations can use Censys to monitor their external attack surface.
- Censys helps identify exposed or vulnerable devices and services.
- Misconfigured devices can pose security risks, contributing to global threats like amplification DDoS attacks.

4. Deprecating SHA-1 Certificates:

- Chrome is deprecating SHA-1 signed certificates due to security concerns.
- Censys is used to assess the prevalence of SHA-1 certificates for HTTPS hosts.
- A significant portion of trusted IPv4 hosts still use SHA-1 certificates.

5. Cipher Suites:

- TLS security relies on strong cipher suites.
- Censys helps analyze the distribution of selected cipher suites by HTTPS hosts.
- Insights include preferences for key exchange methods and encryption algorithms.

Why Censys??

*Censys is a search engine for Internet-connected hosts and certificates which helps information security practitioners discover, monitor, and analyse devices that are accessible from the Internet.

*Censys regularly probe every public IP address and popular domain names, curate and enrich the resulting data, and make it intelligible through an interactive search engine and API.

Who uses it more:

*Enterprises use Censys to understand their network attack surfaces. *CERTs and security researchers use it to discover new threats and assess their global impact.

How it contributes in security domain:

*It plenty of helpful information that could guide some defense forces to obtain more useful findings and decide on how they could tackle some concerns. From this perspective, the Censys is a powerful security tool and at some level – it could serve as a great threat intelligence collector.

*As it's well-known, the majority of threat intelligence could get found below the surface and the Censys itself can grab only data being the part of the visible web. It would use the quite popular Z-map algorithm and it would cope with the wide search only. In other words, this sort of search engine would not go below the internet surface.

Censys follow privacy regulations:

It never attempt to bypass any technical barriers, exploit security problems, or otherwise access non-public-facing services, and it follows community best practices to reduce any burden on remote networks. The only data it receives is information that is publicly visible to anyone who connects to a particular address and port.

Cost or price:

Censys sells paid accounts, but you can also register for a free non-commercial account, which is limited to 250 queries/month. You can view some info even without an account.

Key features:

- 1) One of the key feature is censys search language, its user friendly.

The screenshot shows the Censys search interface. At the top, there's a search bar with 'security' entered. Below the search bar, there's a 'Results' section with a 'Try CensysGPT Beta' link. On the left, there's a 'Host Filters' sidebar with three main categories: 'Labels', 'Autonomous System', and 'Location'. Each category has a list of filters with counts. The 'Labels' category includes 'remote-access' (9.72M), 'login-page' (5.52M), 'jquery' (5.39M), 'network.device' (4.87M), and 'network.device.firewall' (3.36M). The 'Autonomous System' category includes 'AMAZON-02' (1.66M), 'MICROSOFT-CORP-MSN-AS-BLOCK' (1.48M), 'KIXS-AS-KR Korea Telecom' (1.14M), 'DIGITALOCEAN-ASN' (987.01K), and 'CLOUDFLARENET' (916.09K). The 'Location' category includes 'United States' (8.68M) and 'Germany' (3.32M). The main content area shows two host results. The first host is '152.67.19.115', which is a Linux system running on ORACLE-BMC-31898 (31898) in Maharashtra, India. It has several services listed: SSH, CHECKPOINT_TOPOLOGY, UNKNOWN, and HTTP. The second host is '152.67.178.224', which is also a Linux system running on ORACLE-BMC-31898 (31898) in Telangana, India. It has similar services listed. At the bottom right, there's a small orange circular icon with the number '81'.

When you search, you'll see a page with the first 25 search results. Each result shows some data, and you can click links to view more info in Censys.

Along the left side, Censys will display host and service filters, including the following:

- Labels (e.g., remote-access, email, file-sharing, database, network-administration)
- Autonomous System
- Location
- Service Names
- Ports
- Software Vendor
- Software Product

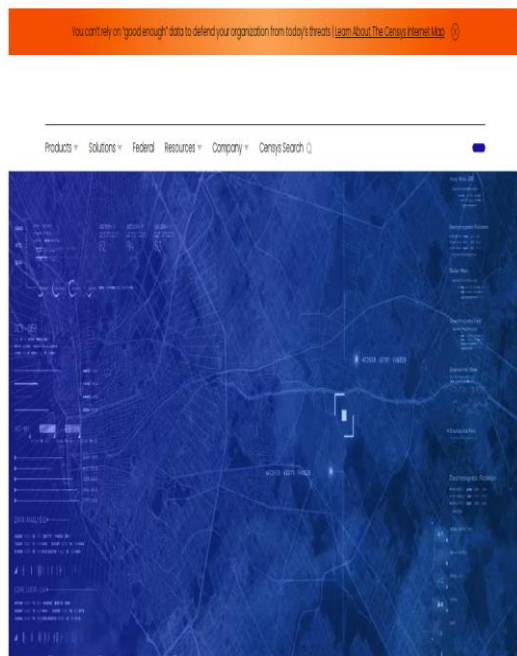
Censys will show up to 5 lines for each of these categories, with links to view more. Near the top of each search results page is also a Report link to build a report.

2) Using Censys for OSINT & CTI investigations:

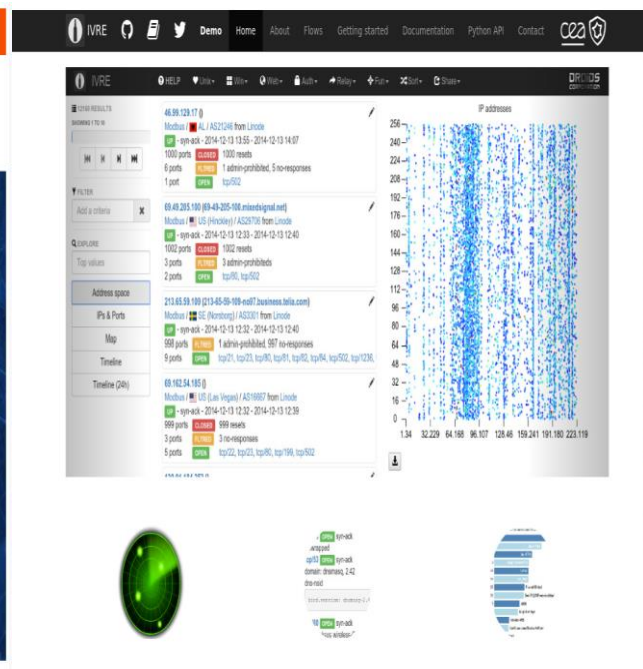
You can use Censys to find the following, which can be useful for OSINT or CTI investigations:

- Hosts with malicious content matching a hash
- History of how a host's attributes have changed over time
- Hosts running software with a particular vulnerability, by searching by CPE-formatted software URI(uniform resource indicator)
- Hosts, certificates, and names connected/related to a particular host, certificate, or name
- Hosts running a specific combination of OS and application
- Hosts with an HTTP service with an open directory list and suspicious file names in their contents

Why Censys is best compared to other search engine (comparing ivre here)??



Censys



Ivre

| Censys | Ivire |
|--|--|
| Censys helps organizations, individuals, and researchers find and monitor every server on the Internet to reduce exposure and improve security. | Ivire Network recon framework, including a web interface to browse Nmap scan results. |
| Censys use wide variety of categories: <ul style="list-style-type: none"> • Attack Surface Management • Web Application Security • Cyber Security • Monitoring Tools • Network Security • Web Security • Security | Ivire uses limited categories: <ul style="list-style-type: none"> • Web Application Security • Attack Surface Management • Security • Monitoring Tools |
| User friendly application. | Has some issues. |

To conclude, censys is an effective search engine and used by majority of people.



Diagram ref: <https://www.saashub.com/compare-censys-vs-ivre?ref=dropdown>

References:

- 1) <https://search.censys.io/search/getting-started>
- 2) <https://jhalderm.com/pub/papers/censys-ccs15.pdf> -- research paper
- 3) <https://search.censys.io/>
- 4) <https://support.censys.io/hc/en-us/articles/360059720271-Common-Host-Query-Example-Searches>
- 5) <https://securityaffairs.com/42725/hacking/censys-search-engine.html>
- 6) <https://www.saashub.com/compare-censys-vs-ivre?ref=dropdown>