CENSYS

HACKER-FRIENDLY SEARCH ENGINE THAT LISTS EVERY INTERNET-CONNECTED DEVICE

An all-new Hacker's Search Engine

KAVYA SRINIVAS MURTHY

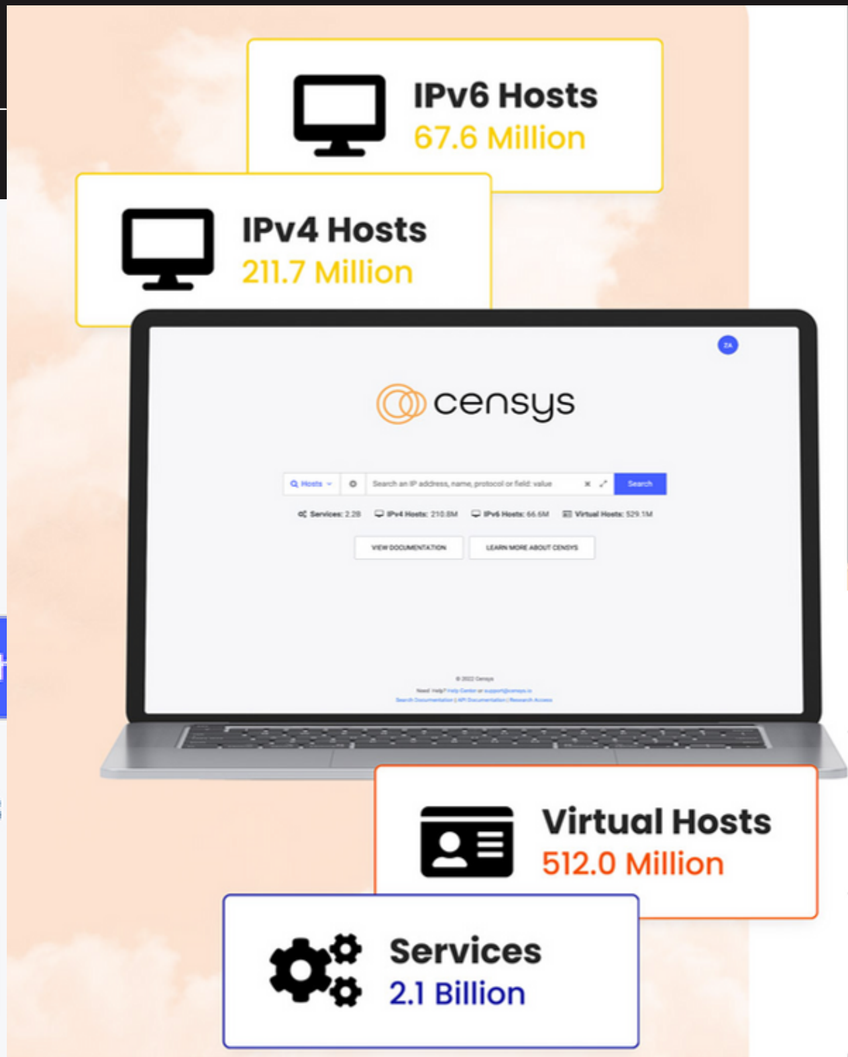
WHAT IS CENSYS?

censys

Hosts | Search an IP address, name, protocol or field: value | Search

Services: 3.5B **IPv4 Hosts:** 246.6M **IPv6 Hosts:** 197.9M **Virtual Hosts:** 1.3B

GETTING STARTED | UPGRADE ACCOUNT | TRY BETA FEATURES



100+ Ports Scanned Daily

Predictive Scanning Across 65K Ports

230M+ IPv4 Hosts

150M+ IPv6 Hosts

800M+ Name-Based Hosts

Data on 3B+ Services Refreshed Daily

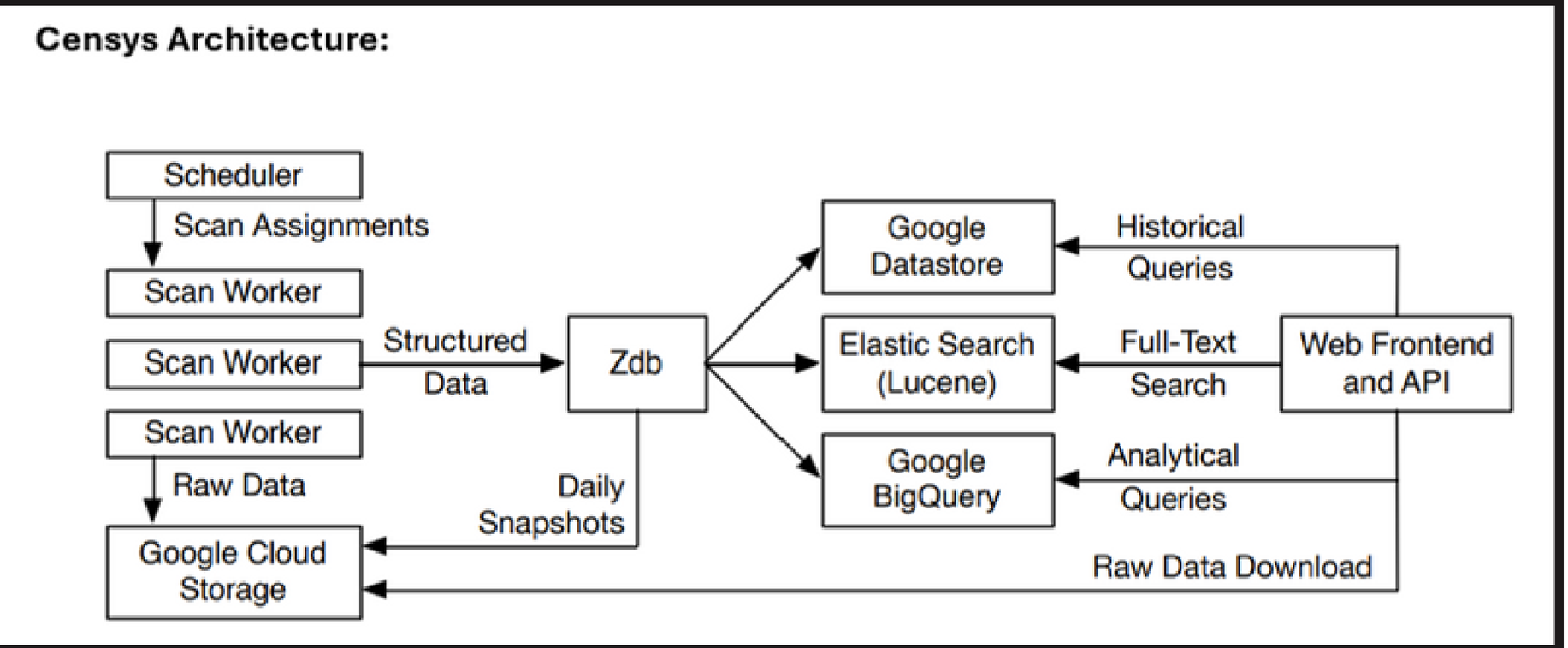
8B+ Certificates

7 years historical data

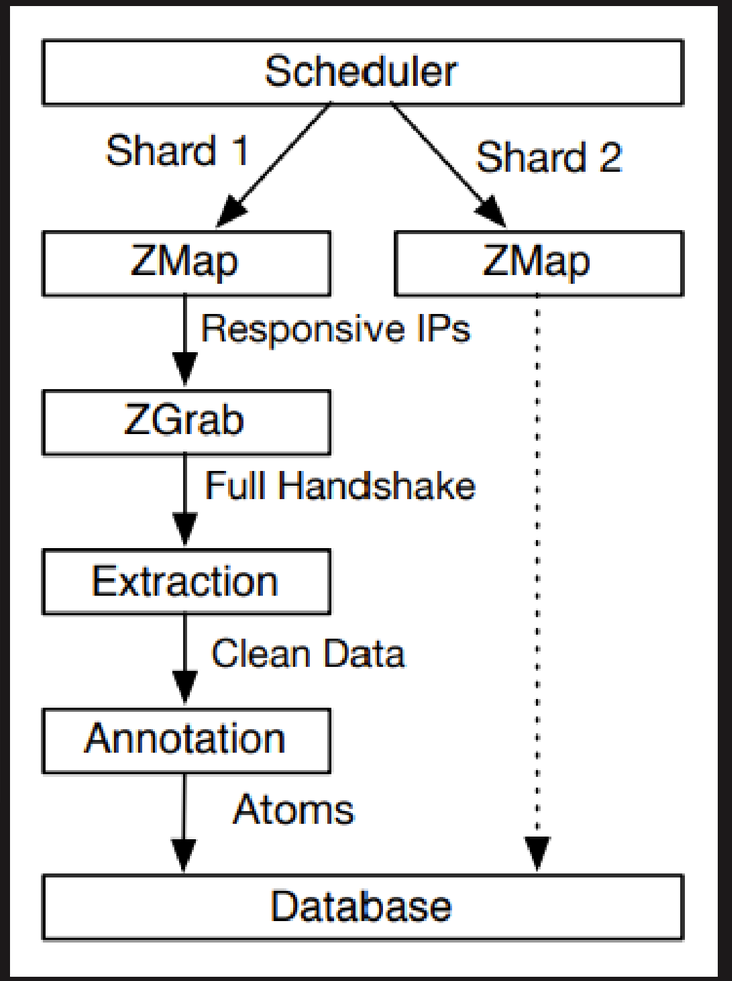
- It’s an all-new Hacker's Search Engine similar to Shodan, which is designed specifically to locate any devices that have been carelessly plugged into the Internet without much attempt at preventing unauthorized access.

- 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.
- 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.
- 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 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.
- 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.
- 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.

CENSYS 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

HOW DOES CENSYS WORK?

1. 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.

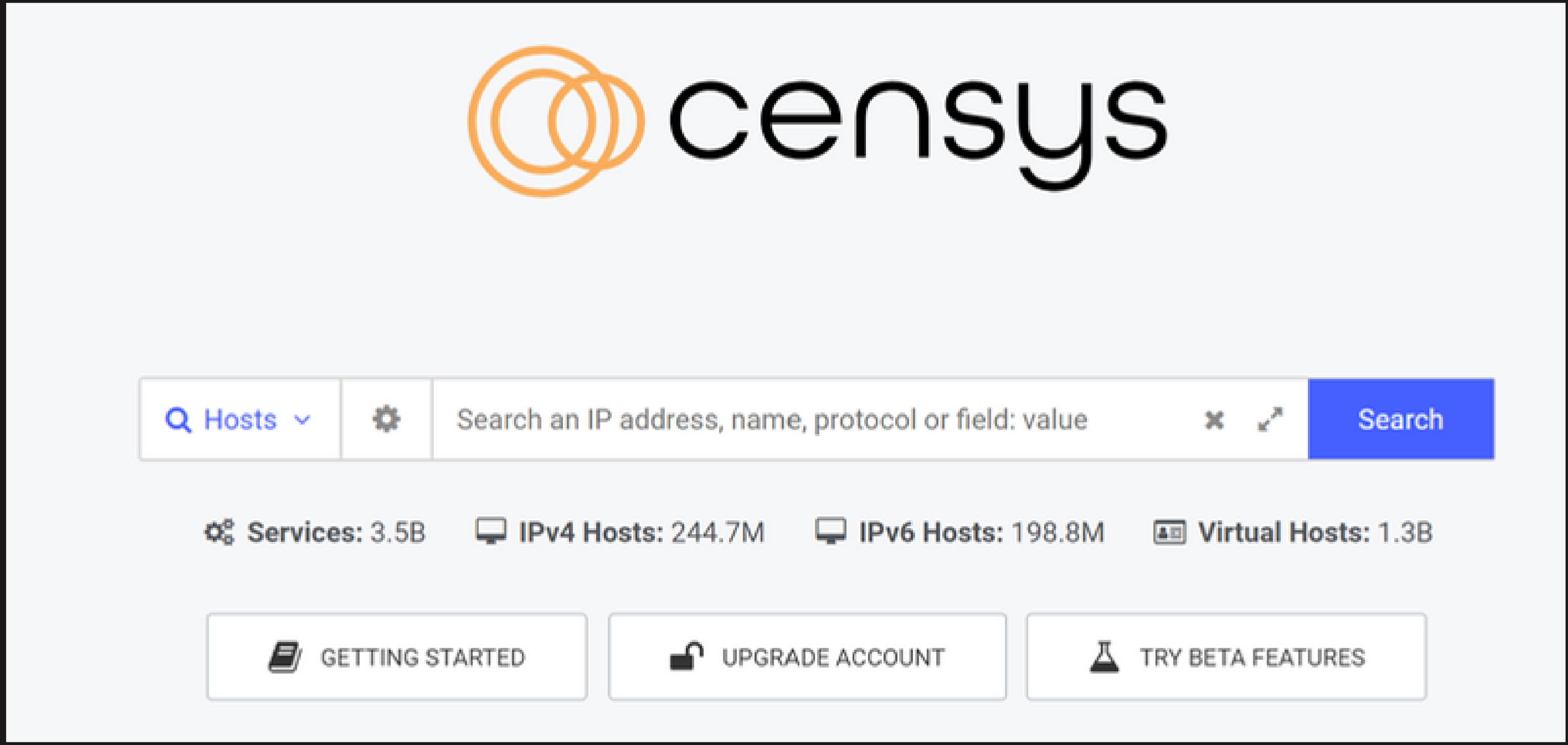
2. This search engine uses two companion tools:
   1) ZMap – an open-source network scanner, the hosts found by ZMap seed pluggable application scanners, which perform a follow-up application layer handshake and produce structured JSON data describing a certain aspect of how a host is configured.
   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.

3. 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.

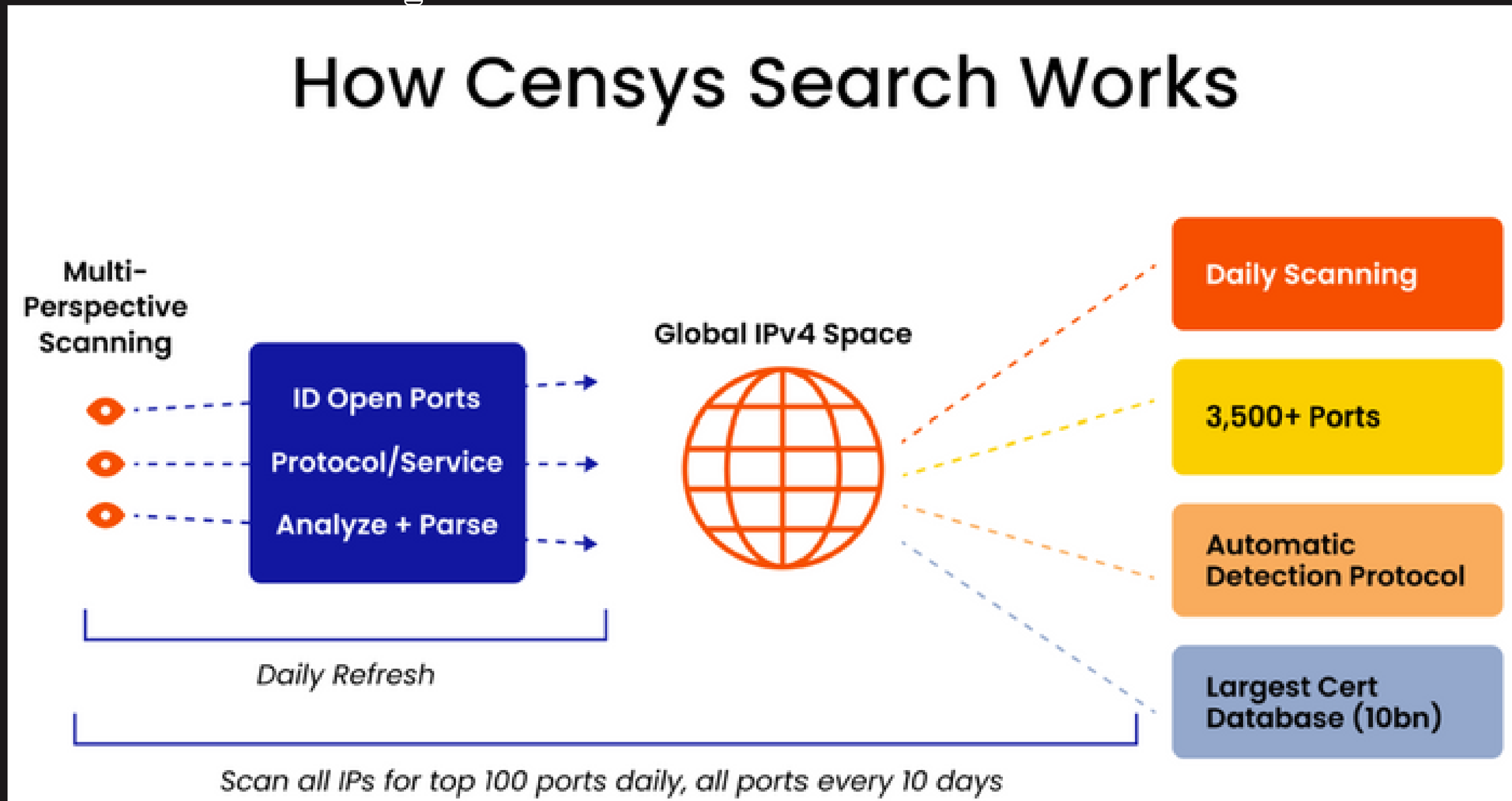
4. ZMap scans over 4 Billion IP addresses on the Internet and collects new data every day. It also helps determine whether the machines on the internet have security vulnerabilities that should be fixed before being exploited by the hackers.

"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.

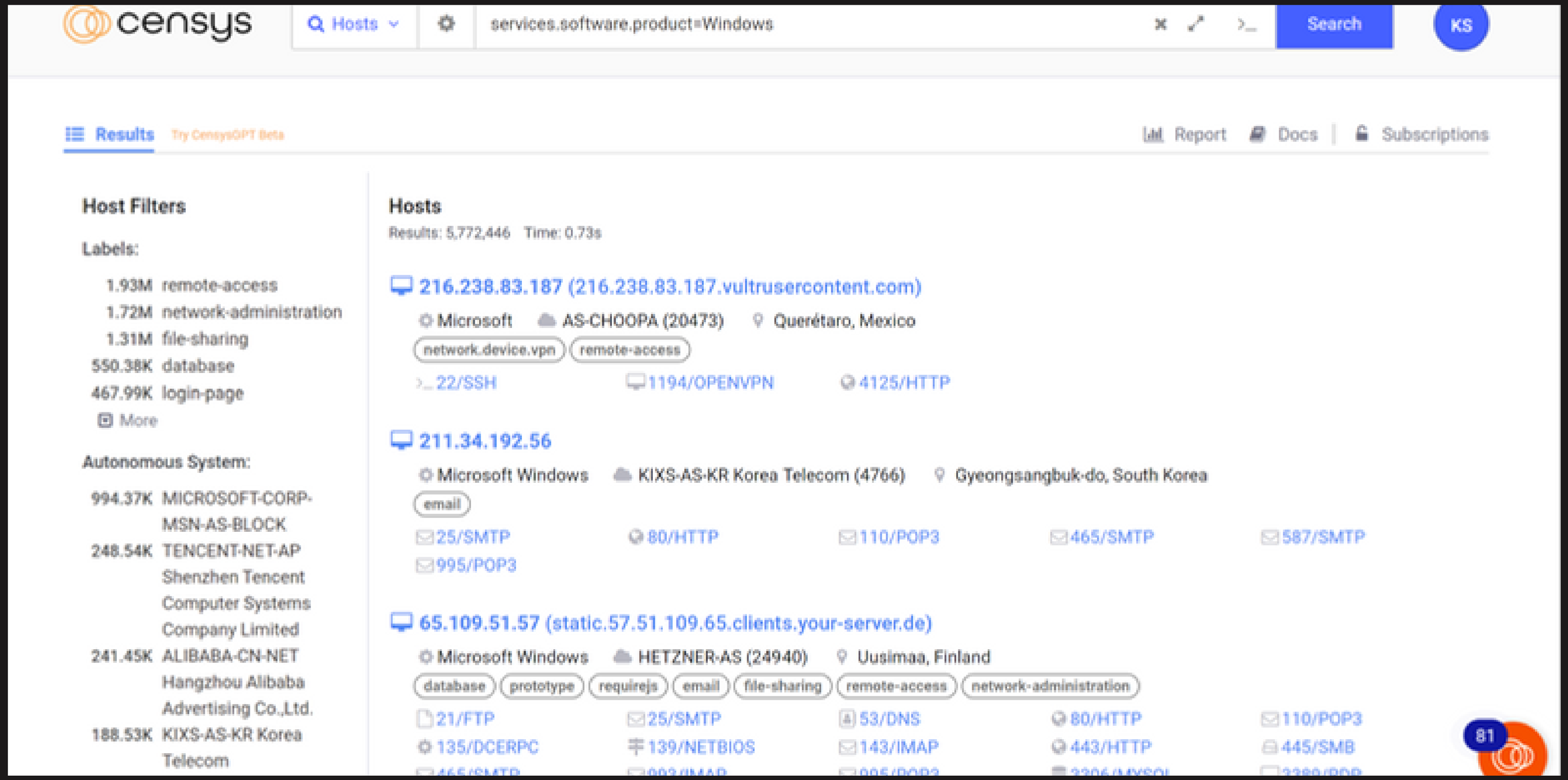
Searches in Censys:



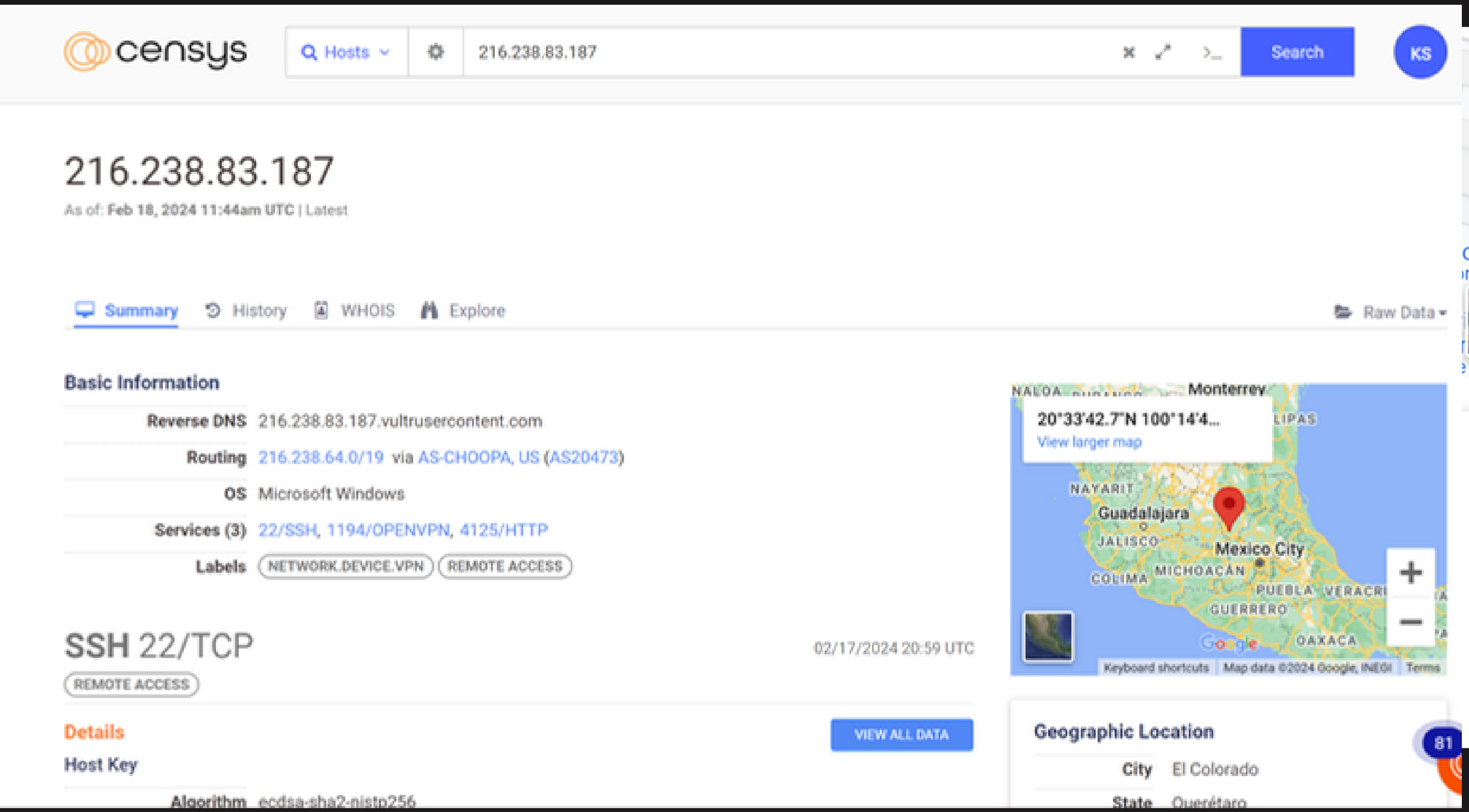
Searches in Censys:



Searches in Censys:



Searches in Censys:



Geographic Location

City	El Colorado
State	Querétaro
Country	Mexico (MX)
Coordinates	20.56185, -100.2452
Timezone	America/Mexico_City

Searches in Censys:

You can click on view definition to view what exactly attributes means.

censys | Hosts | 216.238.83.187 | Search

Host

Attribute	Value
ip	216.238.83.187
location.continent	North America
location.country	Mexico
location.country_code	MX
location.city	El Colorado
location.postal_code	76246
location.timezone	America/Mexico_City
location.province	Querétaro
location.coordinates.latitude	20.56185
location.coordinates.longitude	-100.2452
location_updated_at	2024-02-11T04:58:03.759889476Z

Host Autonomous System

Path	Type	Docs
autonomous_system	object	
autonomous_system.asn	unsigned_long	The ASN (autonomous system number) of the host's autono
autonomous_system.bgp_prefix	ip_range	The autonomous system's CIDR.
autonomous_system.country_code	keyword	The autonomous system's two-letter ISO 3166-1 alpha-2 co RU, ...).
autonomous_system.description	text	Brief description of the autonomous system.
autonomous_system.name	text	The friendly name of the autonomous system.
autonomous_system.organization	text	The name of the organization managning the autonomous s

And view all the tab fields by expanding them, to get overview of host.

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
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.

01 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.

02 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.

03 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.

04 HOST FIELDS

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

05 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.

06 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.

07 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.

08 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.

09 SOFTWARE FIELDS

TLS is a service-agnostic cryptographic pWithin each service object, a software array shows software information in the Common Platform Enumeration (CPE) version 2.3 format.

OVERVIEW OF CENSYS SEARCH LANGUAGE:

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. Also, 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”

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.
services.software.product: Mac

· If you want to search for an exact match replace the colon between the field name and value with an equals sign (=).
services.software.product=Windows

3

WildcardSearches: 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.
location.county: Uni*

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

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

4

Boolean searches: CSL supports logical operators (AND, OR, NOT) and parentheses to combine multiple search criteria effectively, allowing for complex queries to be constructed.
OR -location.country: Canada or location.country: “France”

services.port: 80 and services.service_name: HTTP

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

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

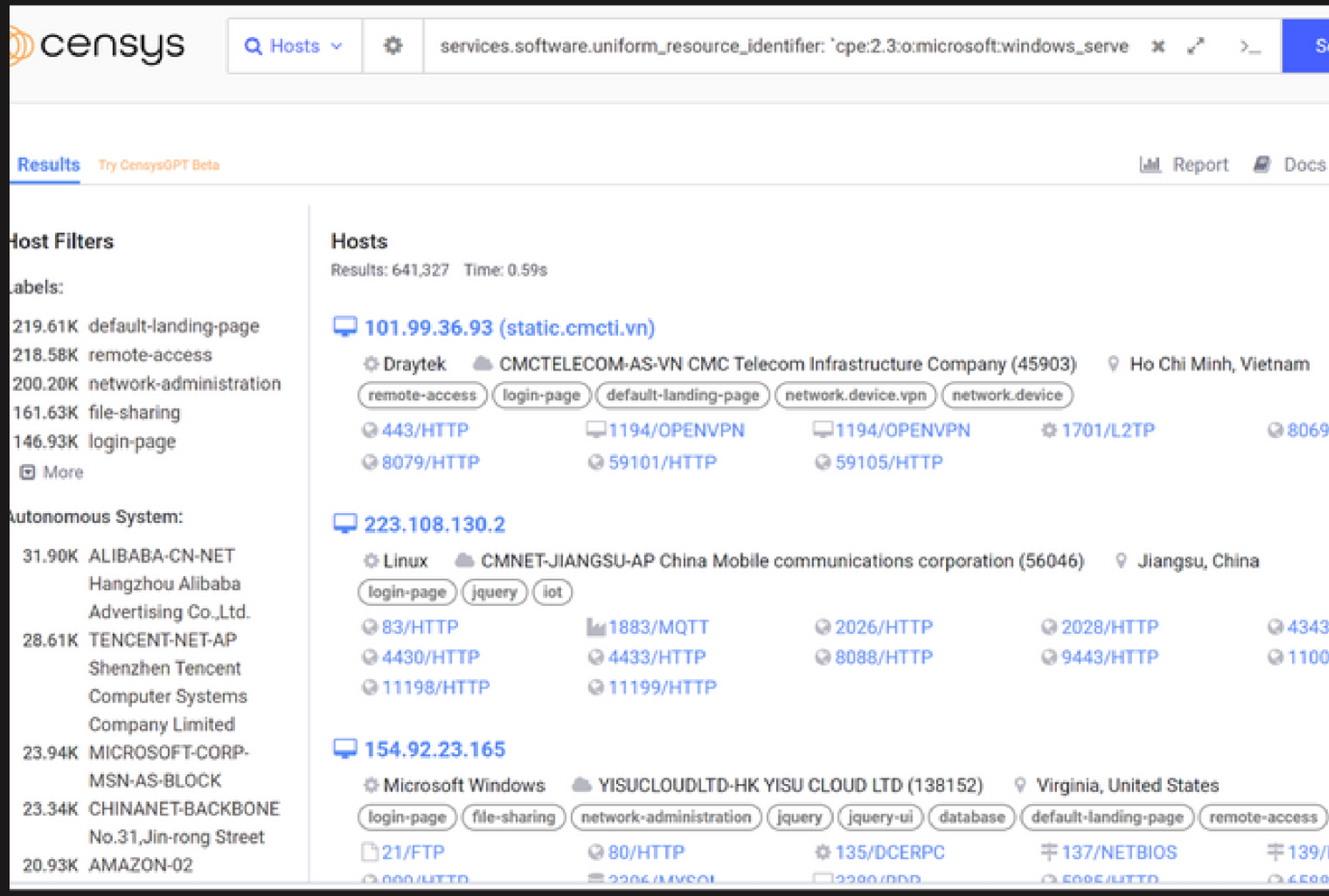
OVERVIEW OF CENSYS SEARCH LANGUAGE:

6

Ranges: Ranges allow users to define a spectrum for numerical values like dates, version numbers, and IP addresses, facilitating searches within specific intervals.
ip: [119.167.243.56 to 119.167.243.201]

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.



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:*:*:*:*:*:*:*:*:*`

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.

Services on Virtual Hosts

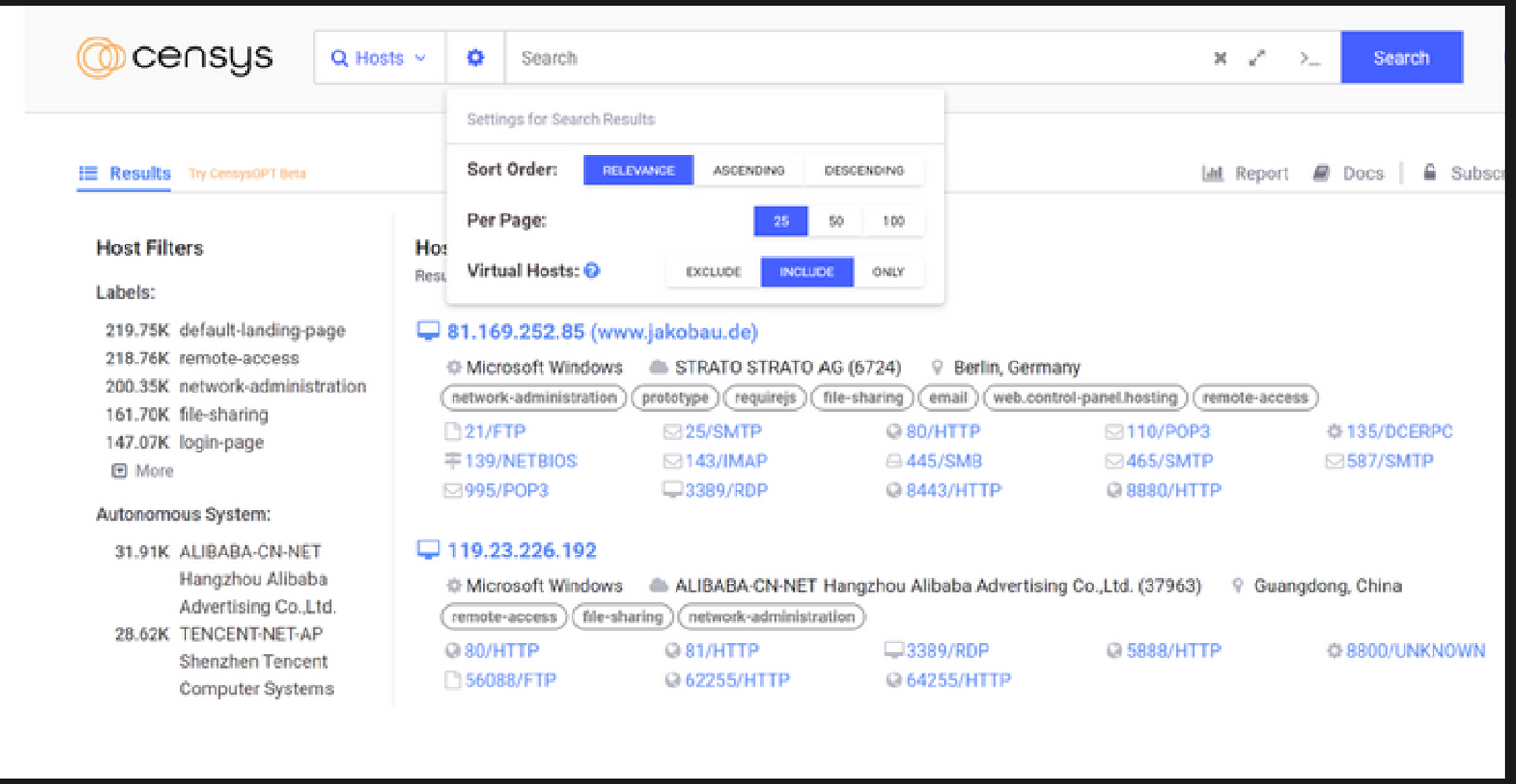
Virtual host records include an array of services (that responded in scan to the virtual host's name) with the same fields as those seen on hosts.
The names of services on virtual hosts are limited compared to those seen on hosts because many protocols do not support name-based differentiation.
Values for services.service_name that can appear on a virtual host:

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

The 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.

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.

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).

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

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"

HONEY POTS IN CENSYS

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.

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.

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.

KEY FEATURES OF CENSYS:

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.

KEY FEATURES OF CENSYS:

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 OF CENSYS:

Key features:

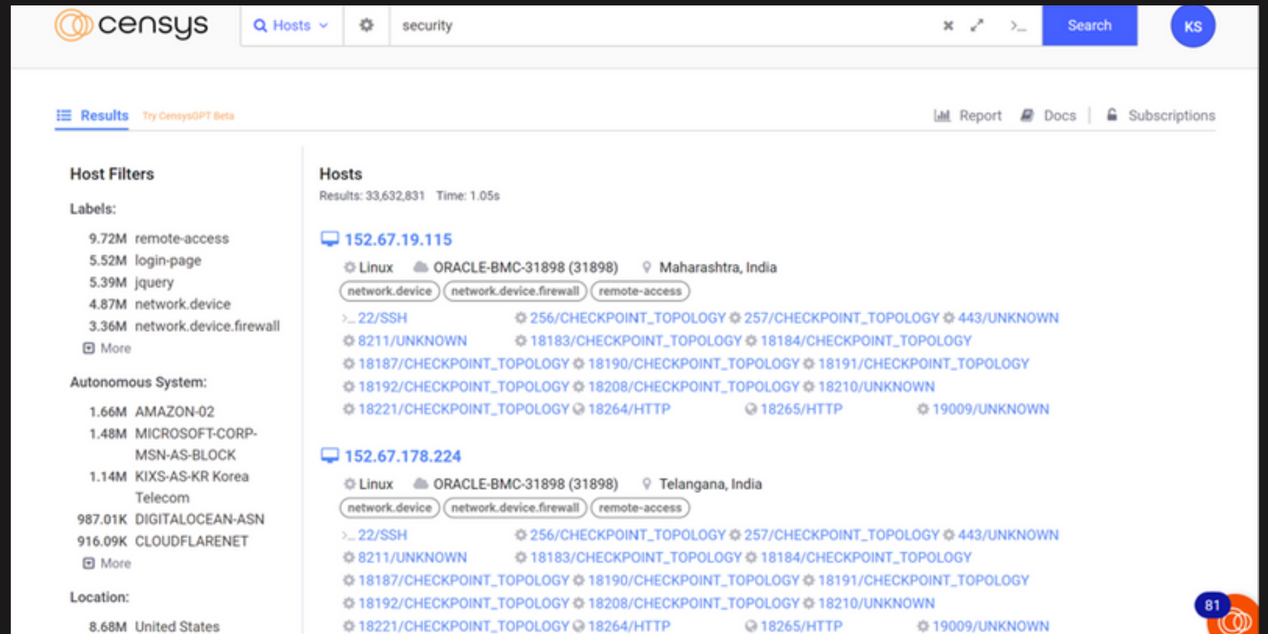
1)One of the key feature is censys search language, its user friendly.

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.



KEY FEATURES OF CENSYS:

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

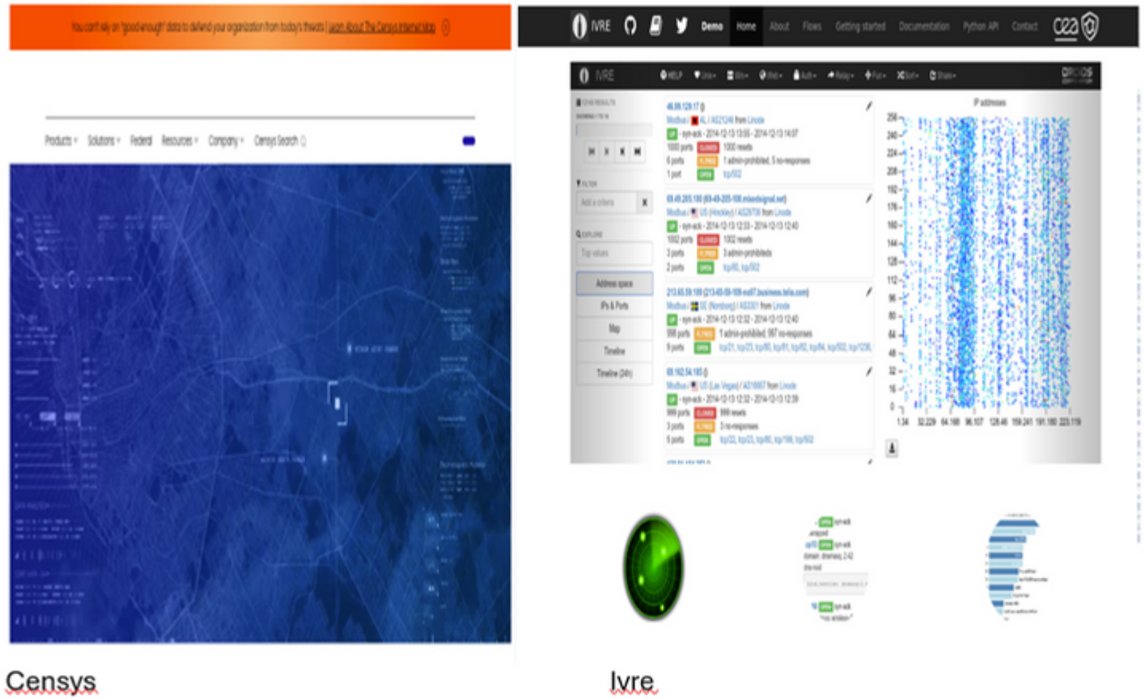
Hosts with an HTTP service with an open directory list and suspicious file names in their contents

KEY FEATURES OF CENSYS:

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

Censys	Ivre
Censys helps organizations, individuals, and researchers find and monitor every server on the Internet to reduce exposure and improve security.	Ivre Network recon framework, including a web interface to browse Nmap scan results.
Censys use wide variety of categories: • Attack Surface Management • Web Application Security • Cyber Security • Monitoring Tools • Network Security • Web Security • Security	Ivre uses limited categories: • Web Application Security • Attack Surface Management • Security • Monitoring Tools
User friendly application.	Has some issues.

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



Censys

Ivre

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

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

THANK YOU!!!