Cloudflare Networking

1.1.1.1

Magic Transit

Speaker info

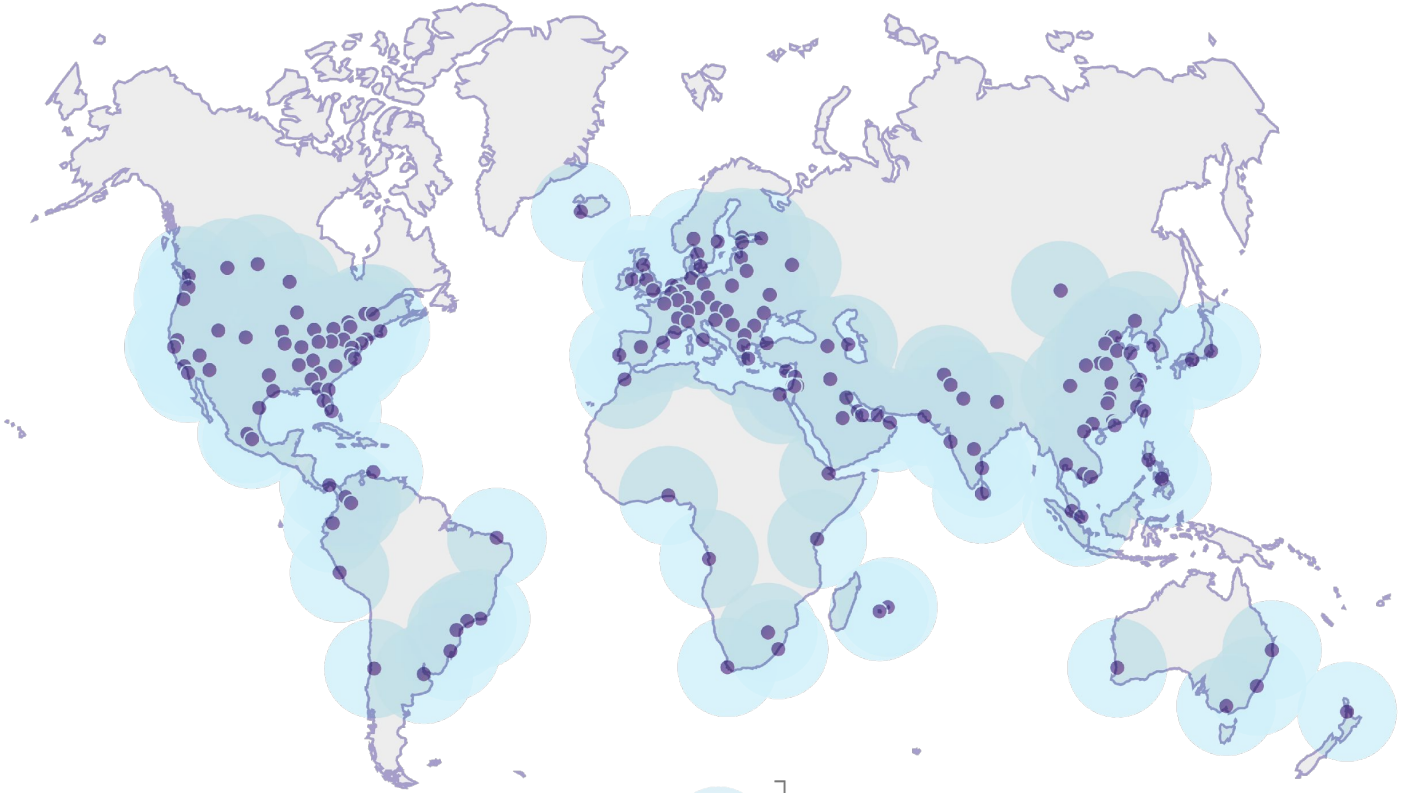
Henry Heinemann

Head of EMEA New Product GTM

hh@cloudflare.com

CLOUDFLARE

Global Anycast Network



Cloudflare city

Approximate area inside which Cloudflare's network is reachable within 100ms via the Internet

CLOUDFLARE Confidential. Copyright © Cloudflare, Inc.

Note: Data as of August, 2019.

20M+
Internet properties

194
Cities and 90 countries

30+ Tbps
Network capacity
DDoS mitigation capacity

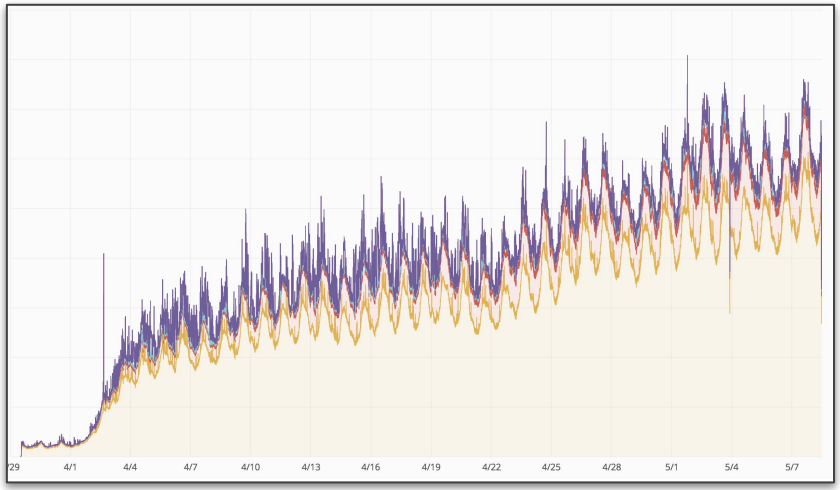
99%
Of the Internet-connected developed world population is located within 100 milliseconds of our network

1.1.1.1

a public resolver focused on privacy

one.one.one.one

global DNS resolver, now at
156 billion queries per day



CLOUDFLARE

1.1.1.1

1.1.1.1
1.0.0.1

2606:4700:4700::1111
2606:4700:4700::1001

setting up the resolver:
https://1.1.1.1

History

- Announced April 1st 2018
- Provided to Cloudflare by APNIC for research and this service
- We knew we would spend a lot of time cleaning up the global Internet to make 1.1.1.1 work

CLOUDFLARE

1.1.1.1

Goals

Speed

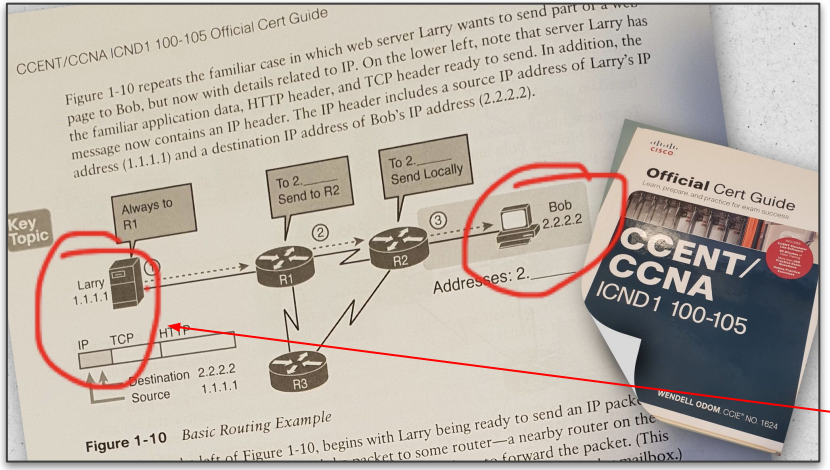
Privacy

Research

Research

Step 32 In the IP Address text box, enter the IP address of the controller's virtual interface.

You should enter a fictitious, unassigned IP address such as 1.1.1.1.



A major hardware vendor

1.1.1.1

Incredibly polluted IP space

Misconception between "assigned" vs "private"

How much noise do you think 1.1.1.1 gets?

CLOUDFLARE

Research

1.1.1.1

#8

Quote:

Originally Posted by **Ulysses_**
Getting tired of typing 192.168. Why doesn't everybody use something simple like 1.1.1.x in a small LAN? What about 0.0.0.x?

I have been using 1.1.1.0/24 subnet for 15+ years on my home LAN and have never found a single instance where any computer in my house ever tried connecting to any address inside the 1.1.1.0-255 range outside my house.

Yes, I realize these are 'publically allocated addresses' but I too got very sick and tired of typing 192.168.blah.blah all the time. I do extensive lab stuff for work where I have servers I build and test in my LAN and am constantly typing IPs all the time.

I still have no regrets about using this subnet. In fact, today in my lab work, I also use 1.1.2.0/24, 1.1.3.0/24, 1.1.4.0/24, 1.1.5.0/24, 1.1.6.0/24, 1.1.7.0/24, 1.1.8.0/24, 1.1.9.0/24 and for the 1.1.2. to 1.1.9. range those are only for lab equipment (have no gateways) for things like iSCSI, vMotion, VSAN and stuff like that so I don't care about them anyway.

You know, if everyone in the world started using 1.1.x.x addresses for home and private LAN use then maybe the industry would change their standard and re-allocate these for official private LAN use, since if someone put a web server on those nobody would ever find their way there. They would be unpopular. Or I guess they are already unpopular because I don't see anyone really using them anyway.

linuxquestions.org/questions/linux-networking-3/why-doesn%27t-everyone-use-1-1-1-x-or-1-1-x-x-or-1-x-x-x-addresses-in-their-lans-4175563056/

CLOUDFLARE

Research

1.1.1.1 polluted space (backbones)

Many backbones seem to have 1.1.1.1 backholed or used - for no real reason

We committed to fixing this by using our measurements to track down, contact and correct these inconsistencies. Here's a partial list of successfully cleaned backbones!

- Airtel, BHTelecom, Beirut-IX, Comcast, Fastweb, ITC, Kazakhtelecom, LG Telecom, Level(3), Liquid Telecom, MTN, Omantel, Rostelecom, SFR, SKBB, Sonatel, STC, Tata, Telecom Italia, Telenor, Telus, Turk Telekom, Turkcell, Voo, XS4ALL, Ziggo
- Many more ...

Thank you backbones. You have helped the Internet improve.

1.1.1.1 polluted space (the edge)

Many CPE routers use 1.1.1.1 for captive portals or configuration screens

- Pace (Arris) 5268
- D-Link DMG-6661
- Technicolor C2100T
- Calix GigaCenter ---- fixed 2018/Jun/12 thanks to a USER
- Nomadix (model(s) unknown)
- Xerox Phaser MFP

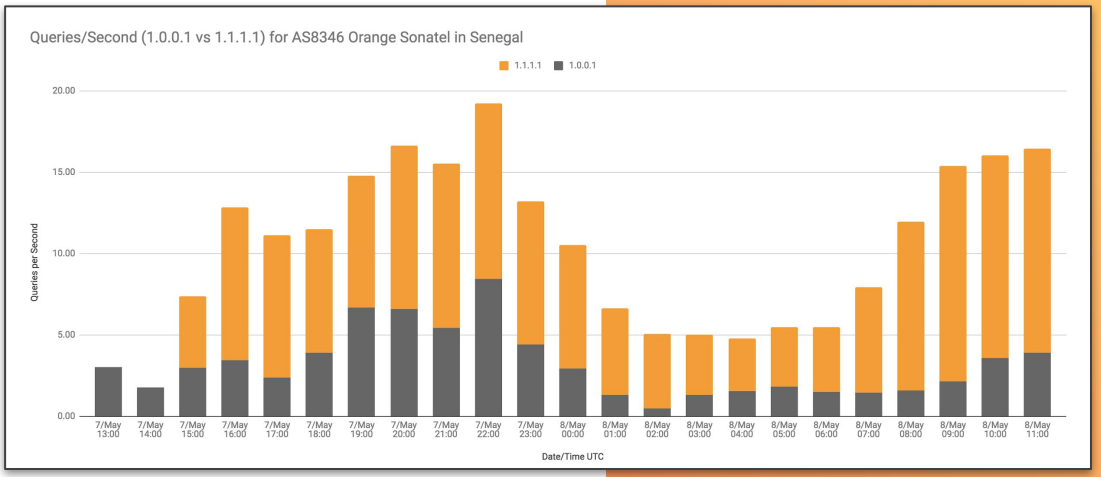
Deployed in the millions globally

1.1.1.1

Fixing networks around the world, one by one...

Why do backbones use this route?

Good question!

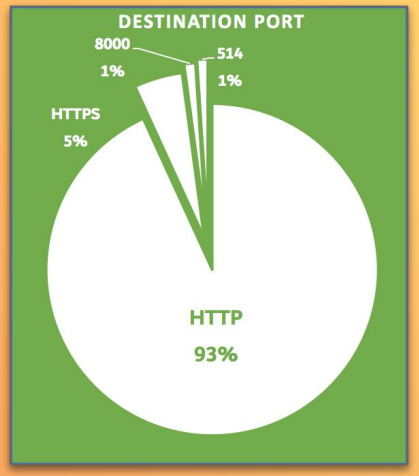


CLOUDFLARE

Research

+10 Gbps of noise!
We estimate legitimate traffic
to be around **7-13%**

1.1.1.1



TCP traffic
80, 443, 8000, 8080, 8090, 8765

UDP traffic
53, 514, 8000, 80, 8090

TP-Link DNS 1.0.0.19

CLOUDFLARE

Privacy

- DNS itself is a 35-year-old protocol
- Never designed with privacy or security in mind.
- It is unencrypted

We focused on privacy:

- Query Minimization *RFC7816*
- Aggressive negative answers *RFC8198*
- No Client Subnet on queries *RFC7871/Experimental*
- DNS-over-TLS (Transport Layer Security) *RFC7858*
- DNS-over-HTTPS protocol DoH (*draft-ietf-doh-dns-over-https*)

CLOUDFLARE

1.1.1.1

We don't store client IP addresses never, ever!

We only use query names to improve DNS resolver performance.

After obfuscation, APNIC research gets access to data.

Cloudflare never stores any information in logs that identifies end user.

All logs are deleted within 24h.

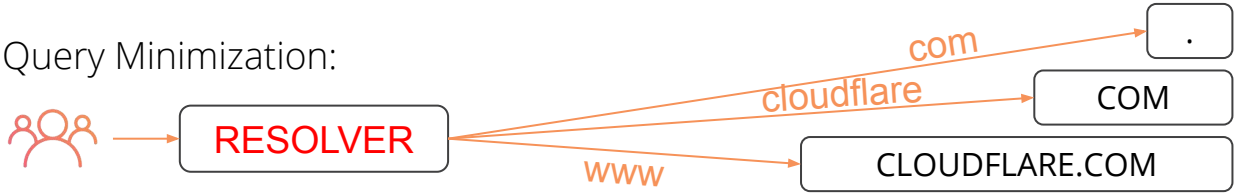
No user data is sold to advertisers or used to target consumers.

Privacy

DNS Query Minimisation

- Resolver can reduce the information leaked to intermediary DNS servers
  - The root, TLDs, and secondary zones
- Resolver only sends just enough of the name for the authority to tell the resolver where to ask the next question.

With Query Minimization:



CLOUDFLARE

1.1.1.1

QNAME contains too much information.

Privacy

Encryption

TLS (Transport Layer Security) is the basis of https encryption.

- DNS-over-TLS (RFC7858) is simply a DNS request wrapped by TLS.
- DNS-over-HTTPS (draft-ietf-doh-dns-over-http) is DNS queries via an HTTPS request.

Resolver, 1.1.1.1 now provides both - at scale!

- Mozilla Trusted Recursive Resolver
  - Cloudflare listed

1.1.1.1

DNSSEC ensures integrity of data between resolver and authoritative server, it doesn't protect privacy of that data!

Specifically, DNSSEC doesn't protect the privacy of the "last mile".

CLOUDFLARE

Speed

Prefill

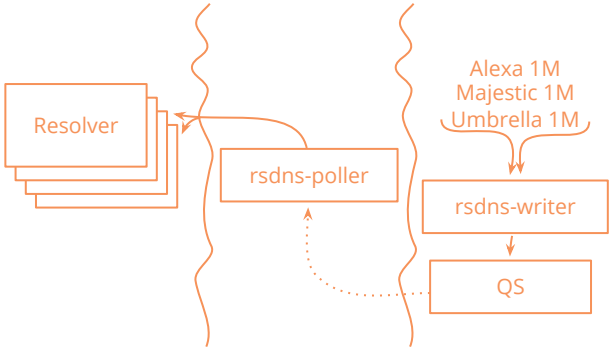
We prefill all caches based on popular domains in a region

- Why: To improve perceived speed and availability
- Popular domains should always be cached
- What is popular?

Backend Multicast

Multicasted cache data across machines within the same data center

- Why: Cache hit ratio goes down with the network size
- Cache hit ratio is everything
- Basically a pub-sub
- Consistent latency



CLOUDFLARE

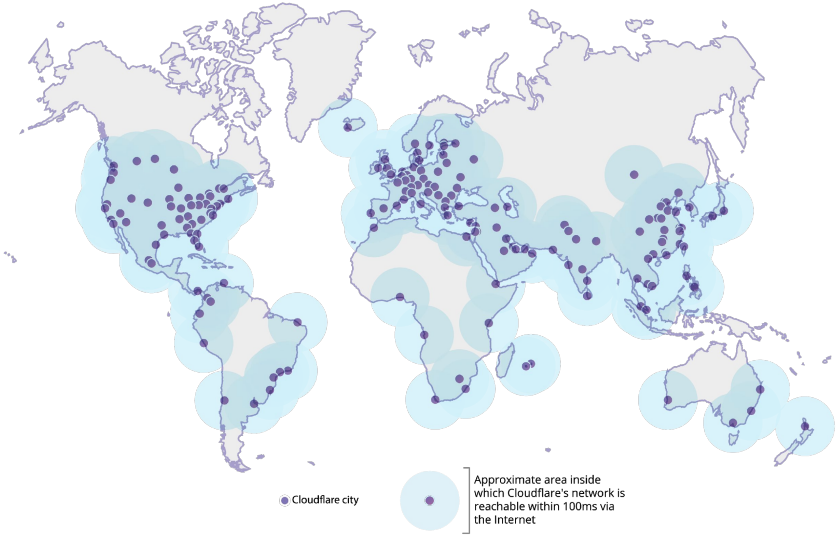
1.1.1.1

DNS name	Query Speed
1.1.1.1	12.65 ms
OpenDNS/Umbrella	19.9 ms
Google	22.78 ms
DNSFilter	23.89 ms
NuSEC	29.21 ms
Neustar	32.51 ms
Norton	34.39 ms
SafeDNS	37.22 ms
Quad9	40.93 ms
Verisign	58.49 ms

dnsperf.com/#!dns-resolvers

Speed

- Our customers' domains are the fastest to resolve on 1.1.1.1 because the DNS answers are on the same servers as 1.1.1.1.
- This is a great reason for websites/mobile apps to use Cloudflare as your DNS provider.



1.1.1.1

DNS name	Query Speed
1.1.1.1	12.65 ms
OpenDNS/Umbrella	19.9 ms
Google	22.78 ms
DNSFilter	23.89 ms
NuSEC	29.21 ms
Neustar	32.51 ms
Norton	34.39 ms
SafeDNS	37.22 ms
Quad9	40.93 ms
Verisign	58.49 ms

global 30 day average

dnsperf.com/#!dns-resolvers

CLOUDFLARE

Summary

- Easy to remember IP addresses
- Support for DOH and DNS over TLS
- Cleaning up routing and CPE devices
- Did I mention it's fast?
- Our mission is to help build a better internet

1.1.1.1

Setting up the resolver:

https://1.1.1.1

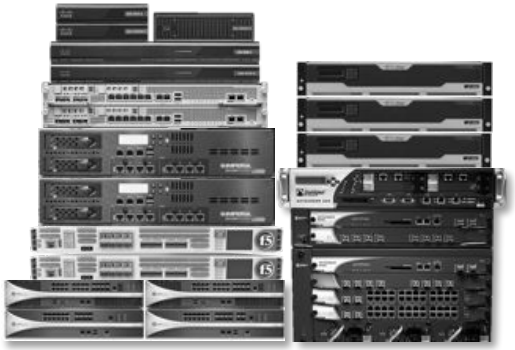
CLOUDFLARE

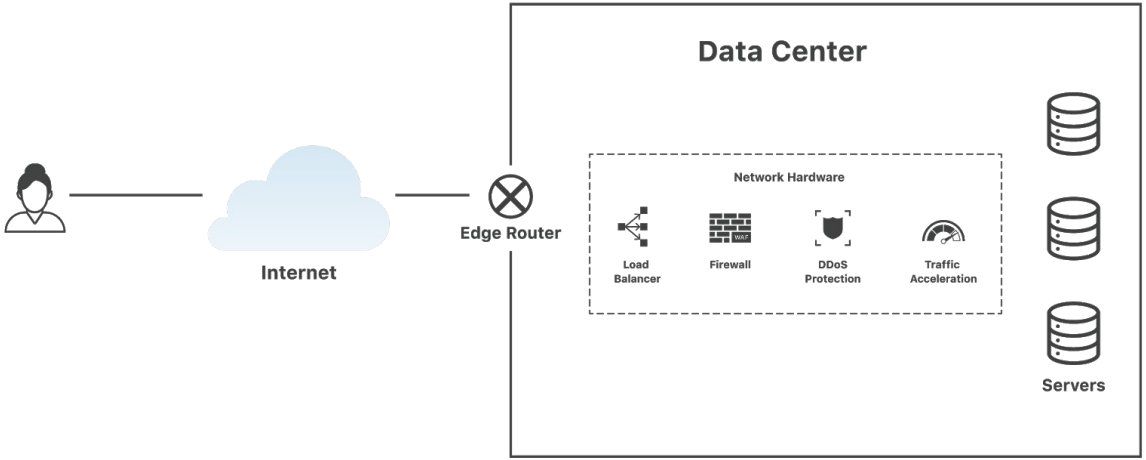
Magic Transit

extending Cloudflare to on-prem networks

Traditional datacenter security architecture

- Costly and complex
- Web of dependencies
- Specialized hardware
- Point solutions





CLOUDFLARE
Confidential. Copyright © Cloudflare, Inc.

Digital Transformation

and the disconnect

Compute, storage, applications are moving to the cloud.

But perimeter networks still remain on-prem.

It’s time for a change.

CLOUDFLARE

Confidential. Copyright © Cloudflare, Inc.

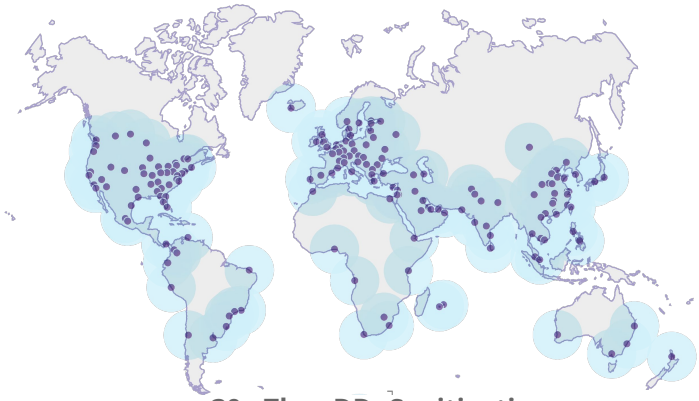
What's needed?

> The **transition** to a global cloud platform that delivers **security**, **performance**, and **reliability** virtual network functions that can be deployed as services on the fly.

CLOUDFLARE

Confidential. Copyright © Cloudflare, Inc.

Our battle-tested network stack, in front of your data center



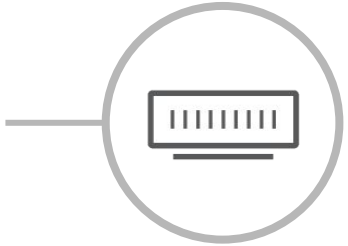
30+ Tbps DDoS mitigation capacity

DDoS protection
30 Tbps network capacity
Near-instant TTM

Network firewall
Granular Allow/Deny rules for IP ranges

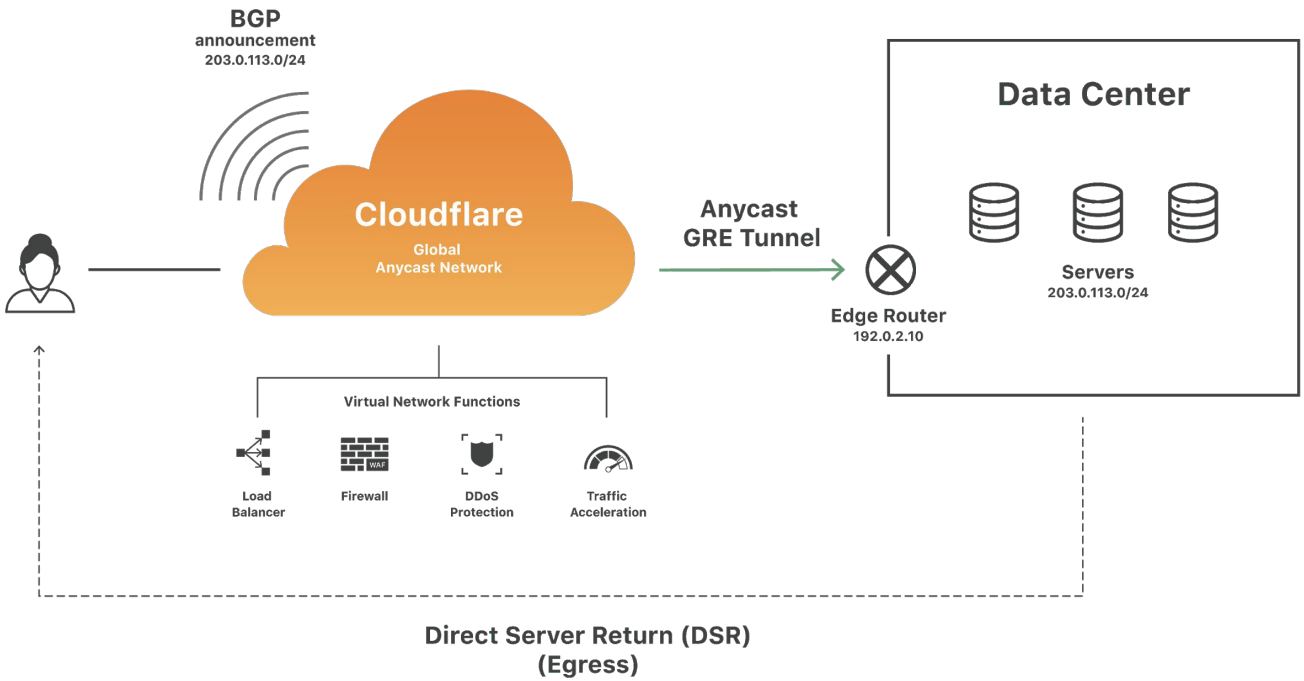
LAYER 3 - IP
(MAGIC TRANSIT)

CLOUDFLARE DATA CENTER



CUSTOMER DATA CENTER

How it works



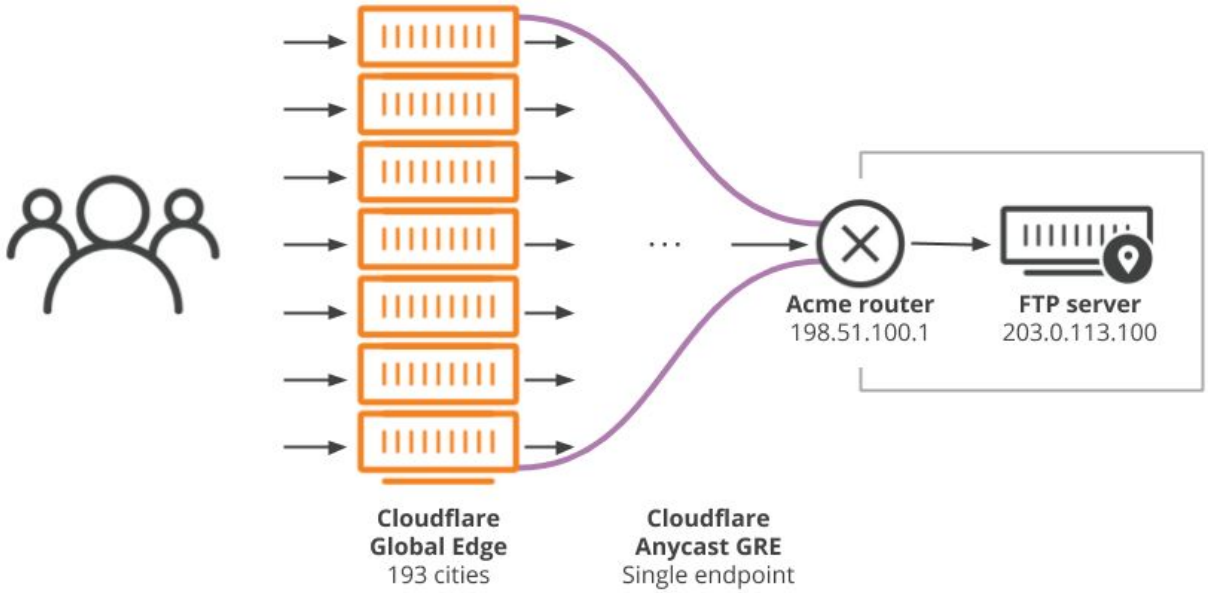
CLOUDFLARE

Connect

Protect & Process

Accelerate

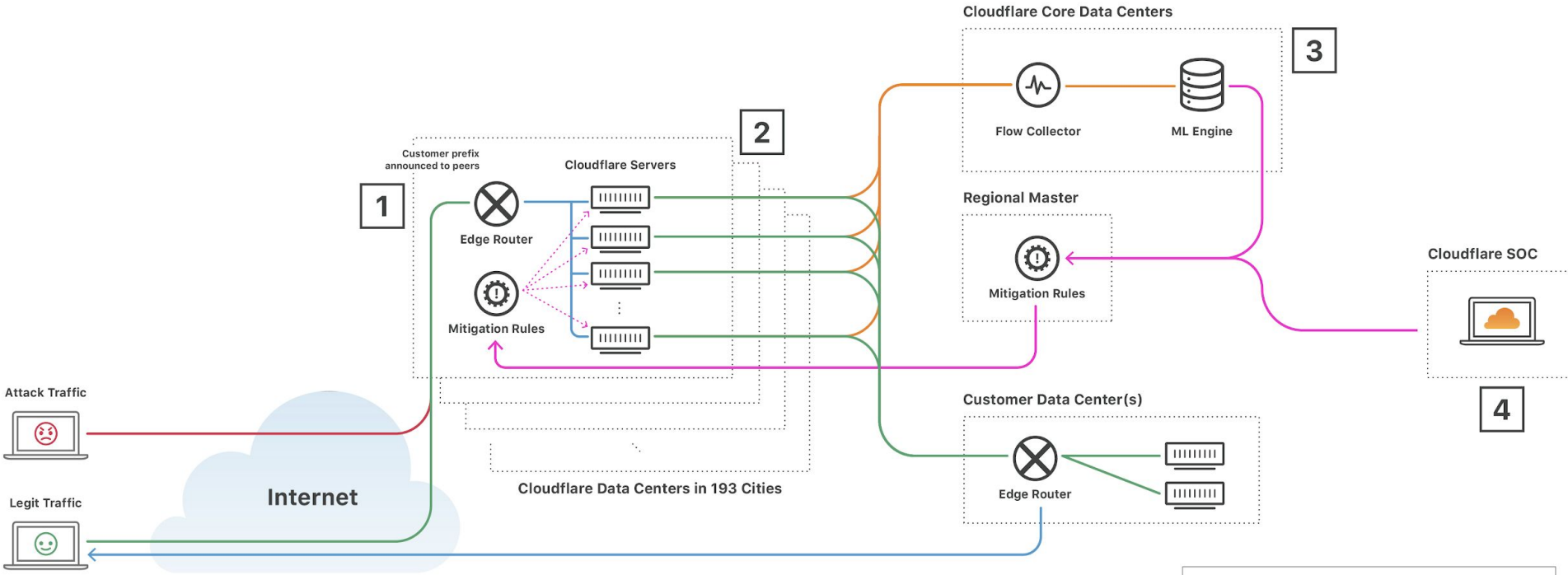
Isn’t a tunnel point to point?



1. GRE is stateless
2. packets are processed independently
3. the tunnel is bound to an *IP address*, not a *specific device*

CLOUDFLARE

DDoS mitigation methodology



1) Static rules (traffic you never want to see)
Mean time to respond: 0 seconds

2) Edge mitigations
MTTR: < 10 seconds

3) Core analyzed/edge enforced mitigations
MTTR: < 5 minutes

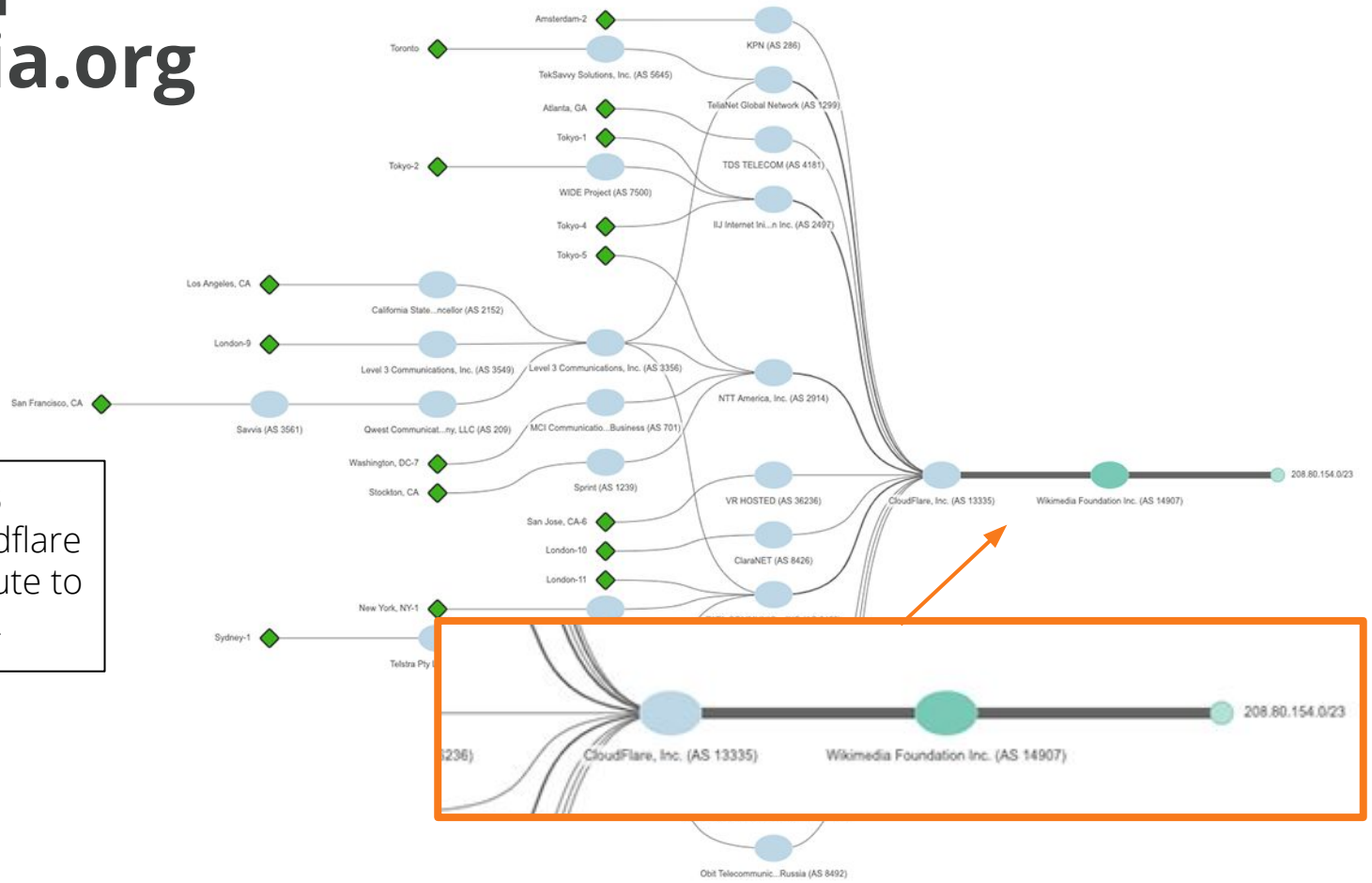
4) SOC in the loop/paging escalation
MTTR: < 15 minutes

- Higher-cost (compute time) mitigations pushed from core, with lower cost rules computed within each colo
- GRE– encapsulated clean traffic
- sFlow samples
- Direct Server Return (DSR) of legitimate traffic

CLOUDFLARE

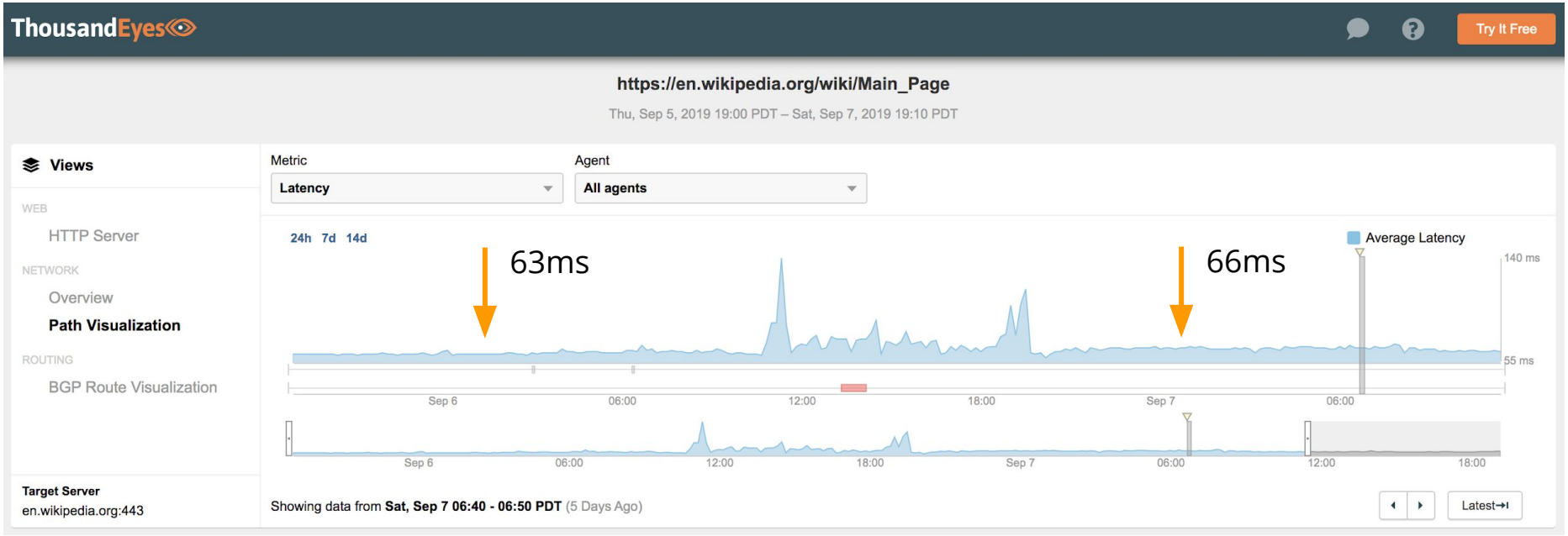
An example: **wikipedia.org**

208.80.154.0/23
is advertised by Cloudflare
as a more specific route to
208.80.152.0/22



CLOUDFLARE

An example:
wikipedia.org



CLOUDFLARE

Summary

L3 + L4 + L7 mitigation
All in the same “cloud”

Stack L4/L7 services on top of Magic Transit

Get rid of network perimeter hardware

Drive down TCO by swapping CAPEX for OPEX

GRE over Anycast

194 globally distributed scrubbing centers *at the same time*

CLOUDFLARE

THANK YOU

For more details, contact me at:

hh@cloudflare.com
linkedin.com/in/henryheinemann

CLOUDFLARE
Confidential. Copyright © Cloudflare, Inc.