Measurement and Analysis of Private Key Sharing in the HTTPS Ecosystem

Frank Cangialosi, Taejoong Chung, David Choffnes, Dave Levin,
Bruce M. Maggs, Alan Mislove, Christo Wilson

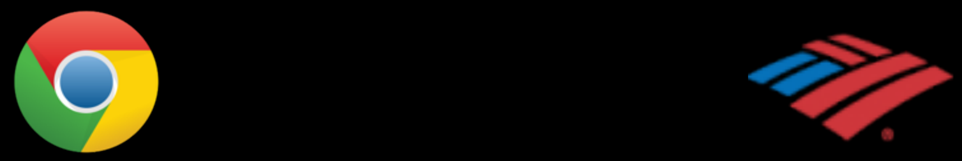
CSAIL

DUKE COMPUTER SCIENCE

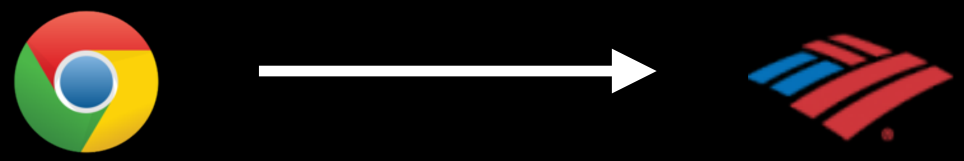
NORTHEASTERN UNIVERSITY BOSTON, MASSACHUSETTS FOUNDED AD 1898

UNIVERSITY OF MARYLAND 1856

How do we know with whom we are communicating?



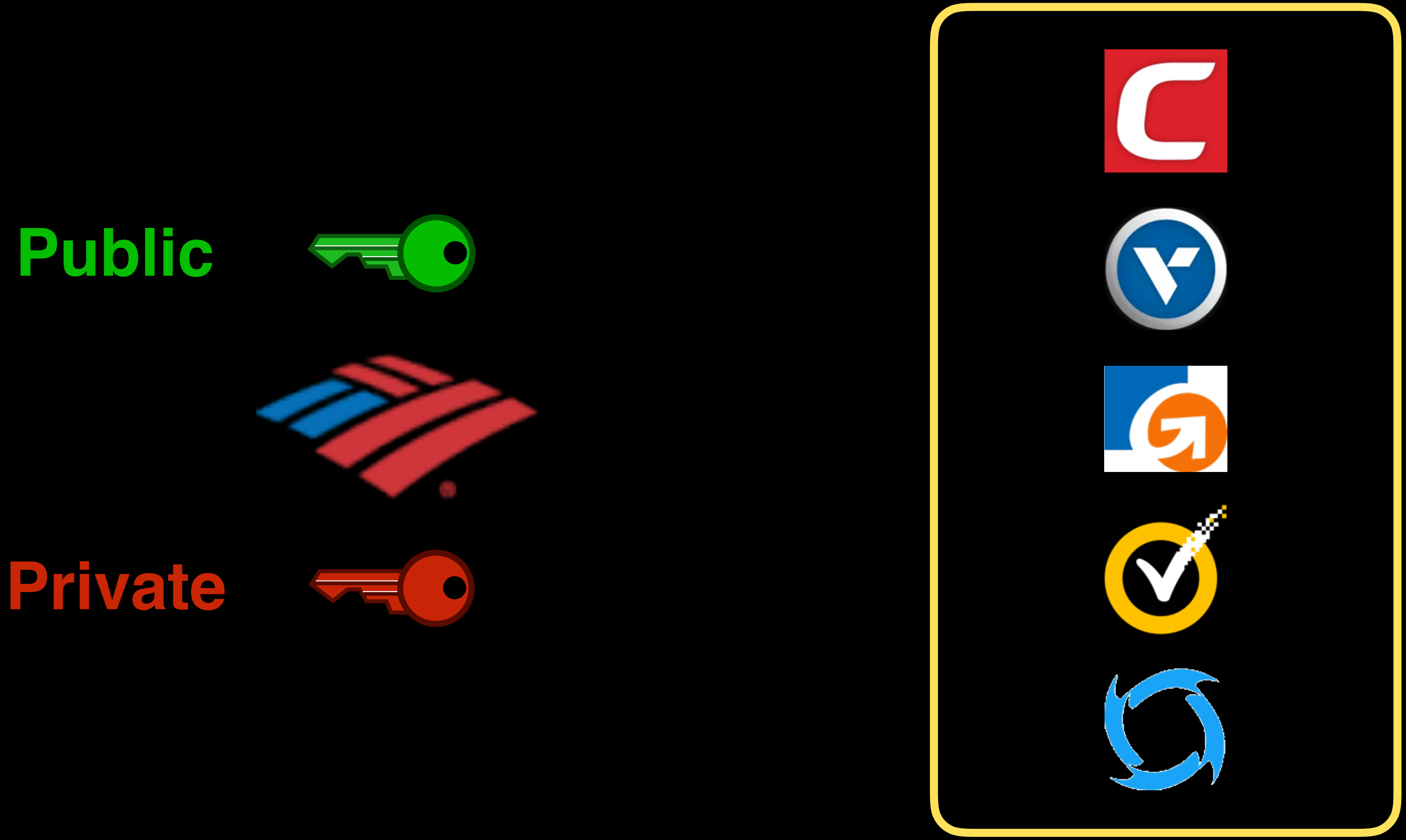
How do we know with whom we are communicating?



Public

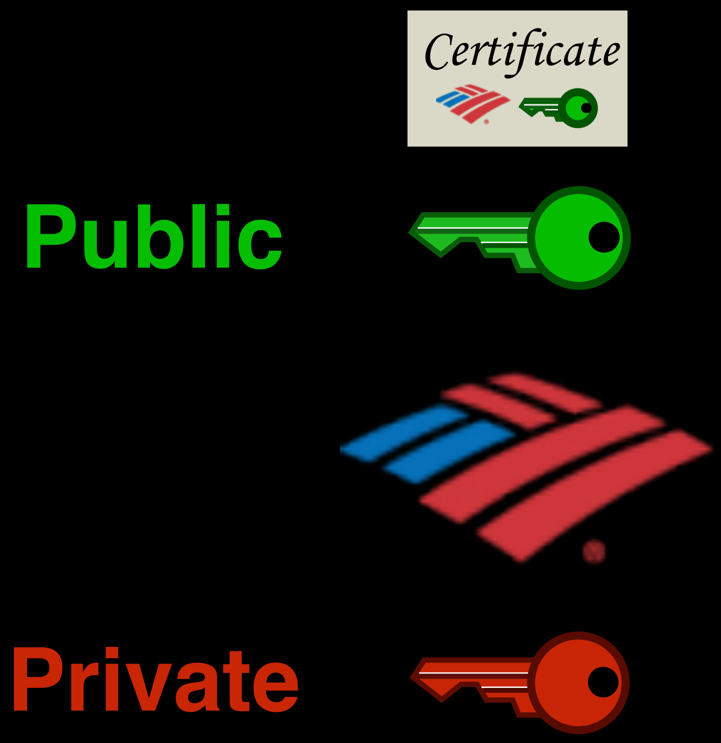
Private

Certificate Authorities



Public

Private

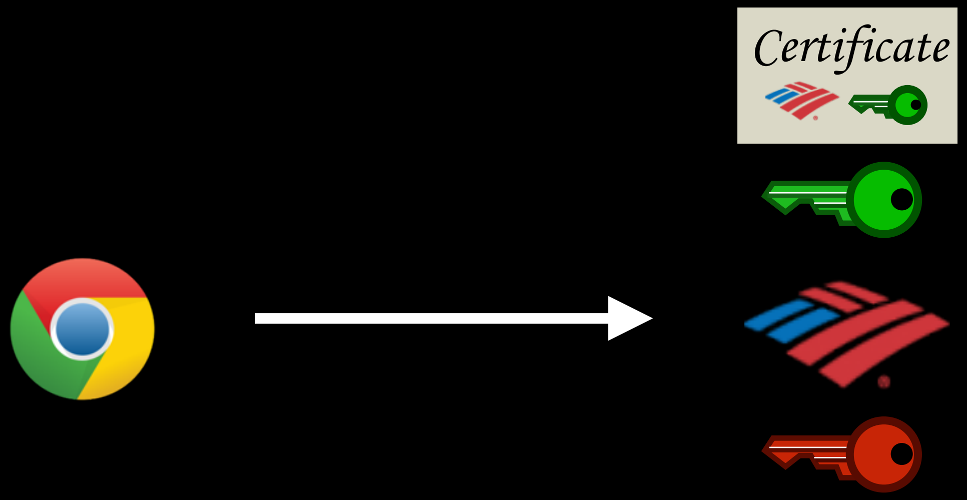


Public

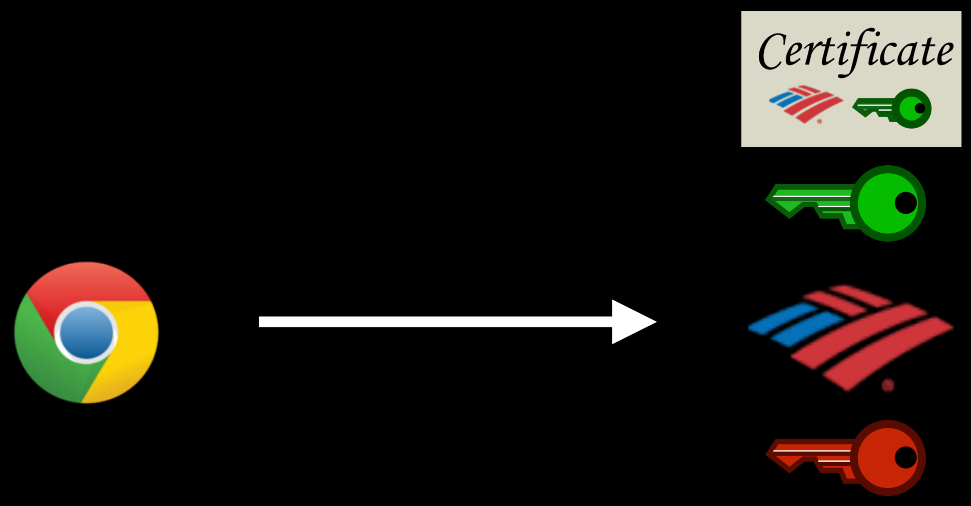
Private

Certificate Authorities

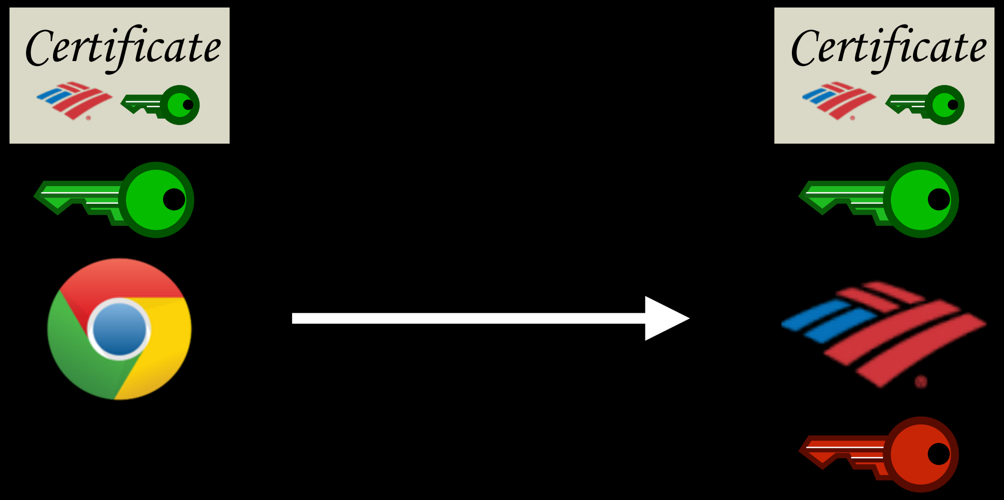
How do we know with whom we are communicating?



How do we know with whom we are communicating?



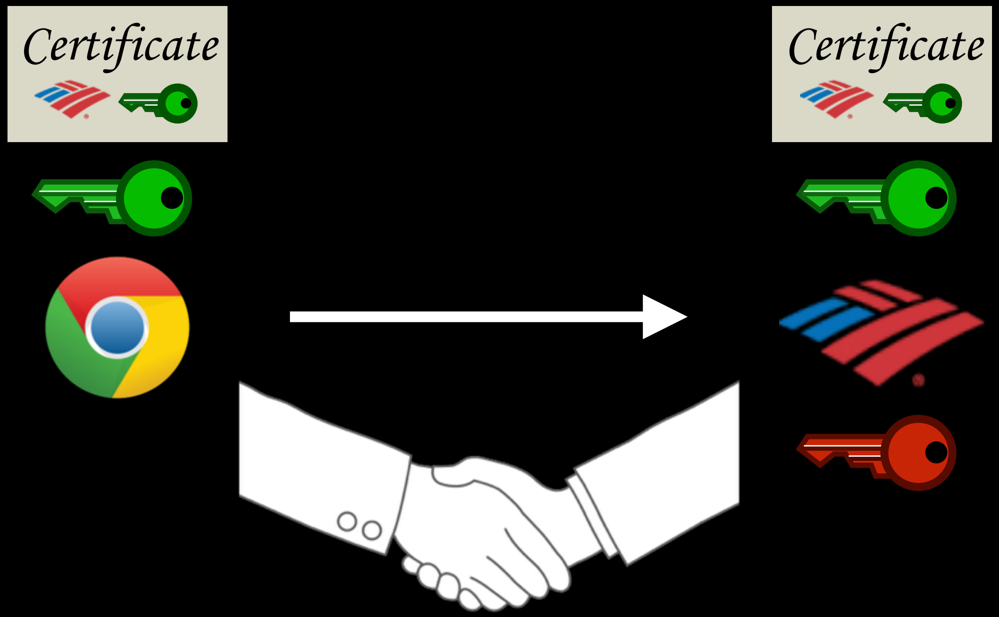
How do we know with whom we are communicating?



Certificate

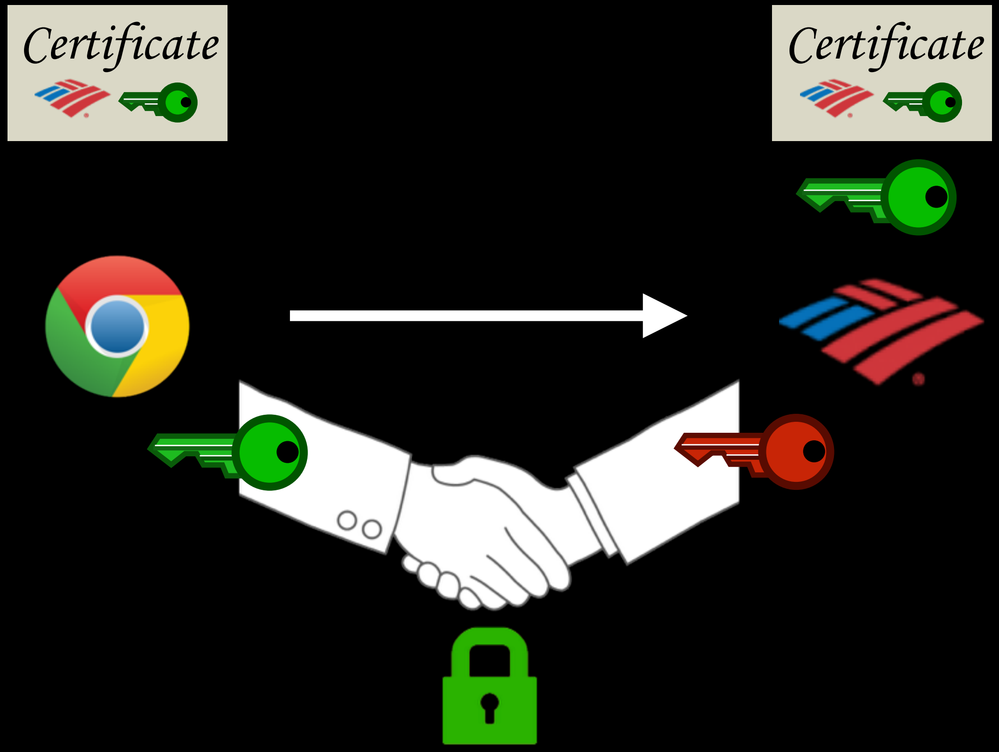
Certificate

How do we know with whom we are communicating?

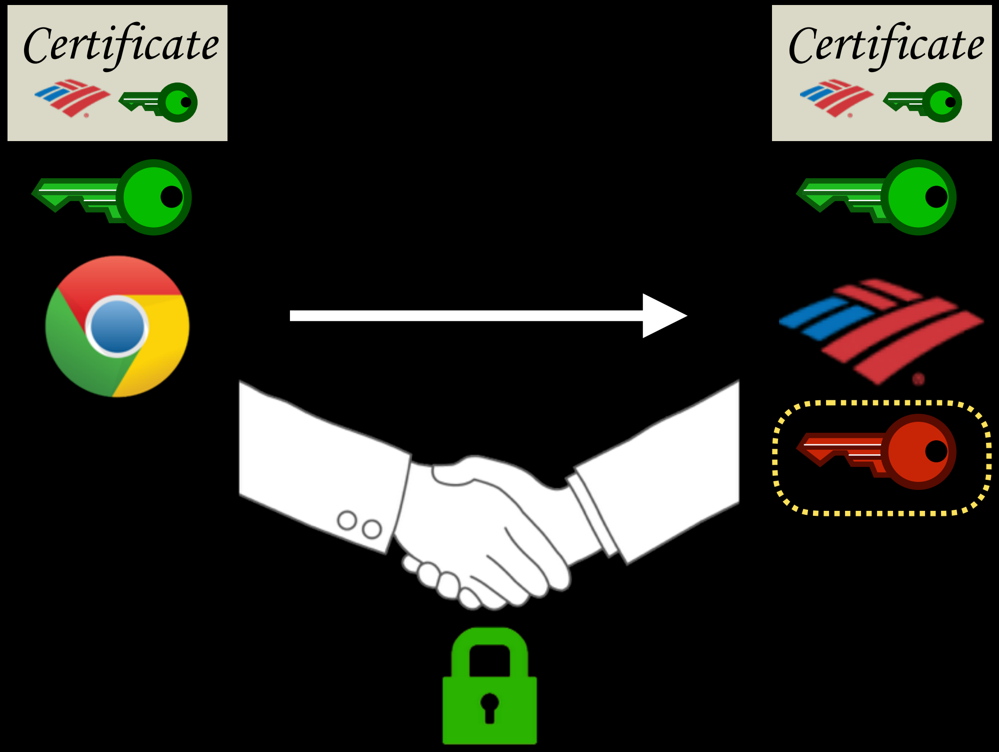


TLS Handshake

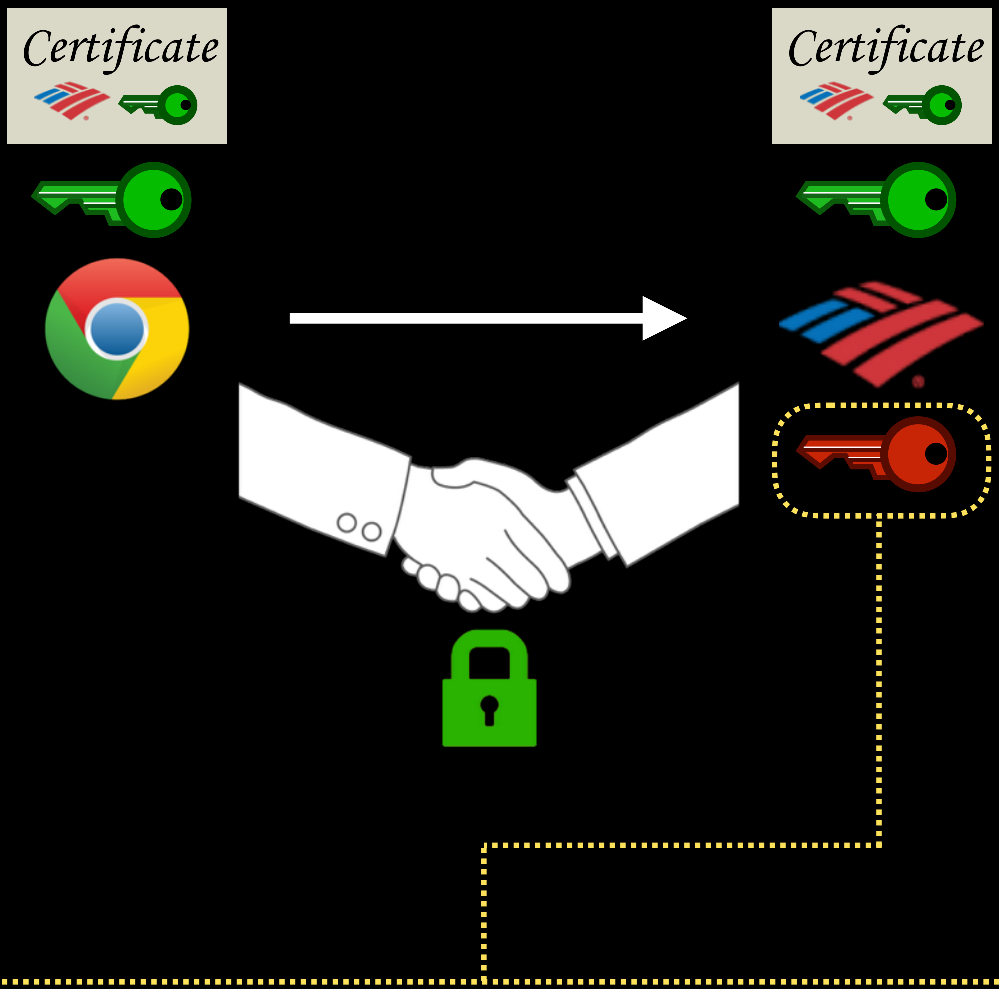
How do we know with whom we are communicating?



How do we know with whom we are communicating?



How do we know with whom we are communicating?



Authentication **fundamentally** assumes:

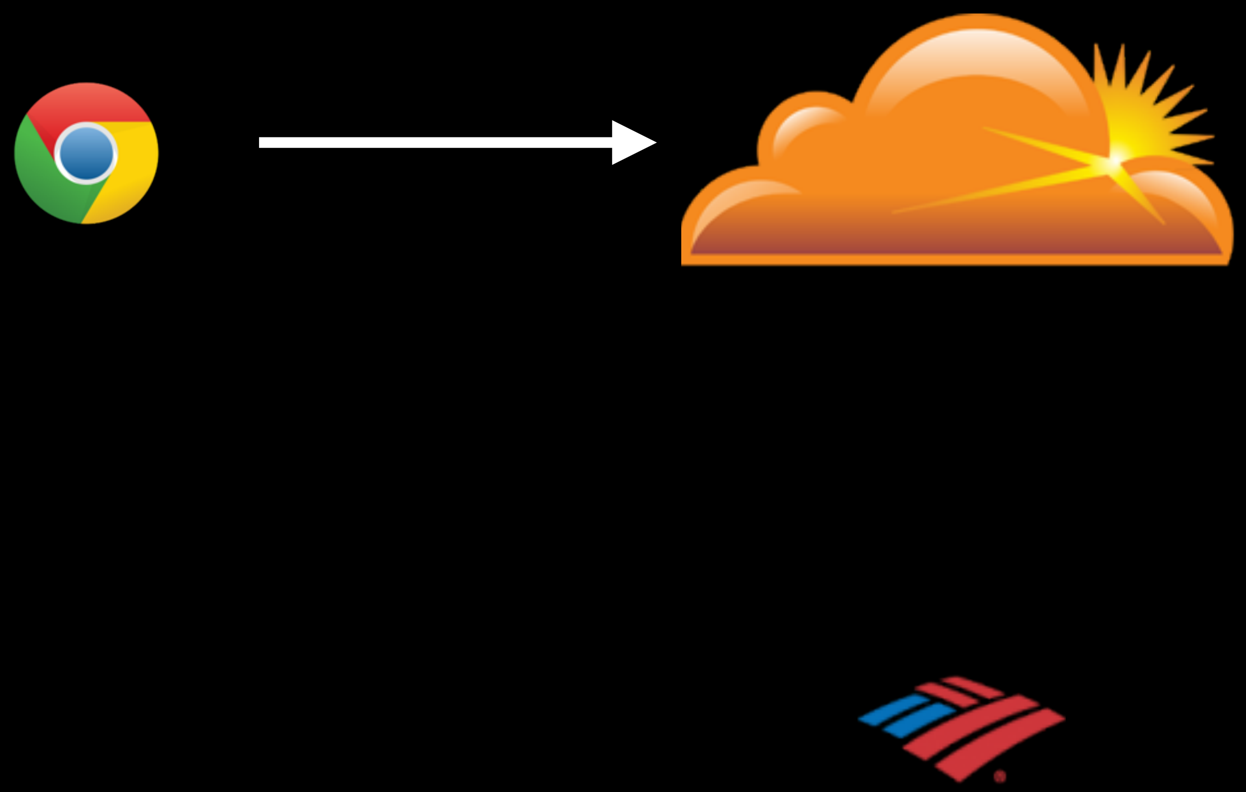
Only **knows**

The PKI in today's web

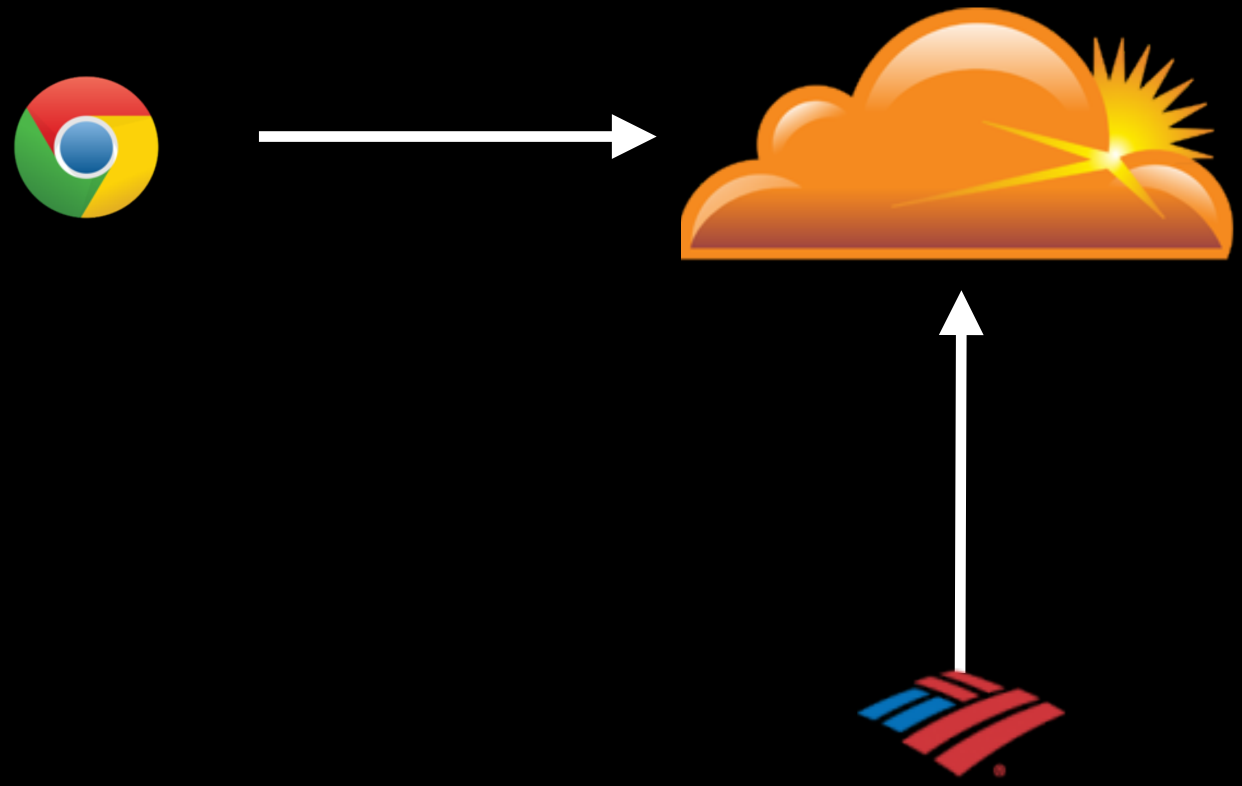
The PKI in today’s web

Rare!

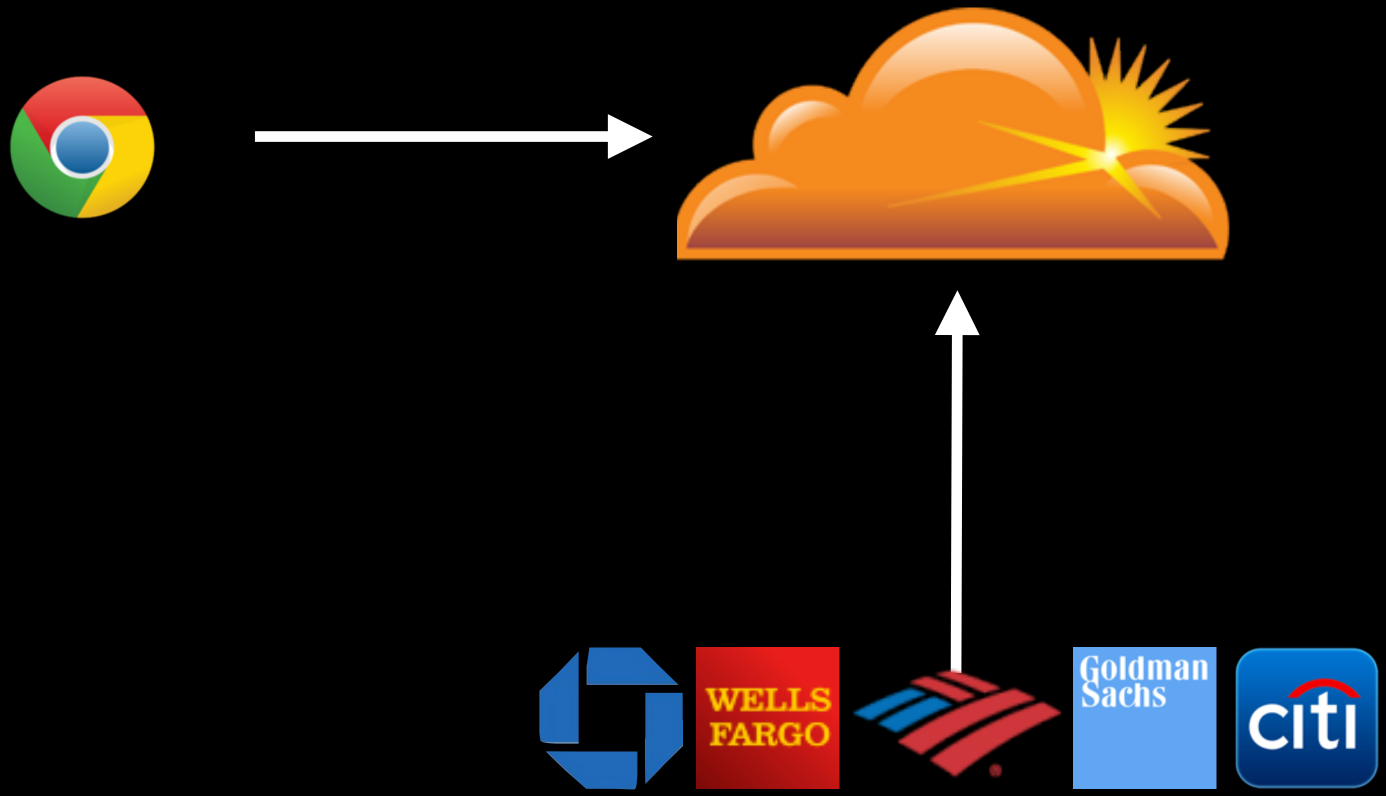
The PKI in today’s web



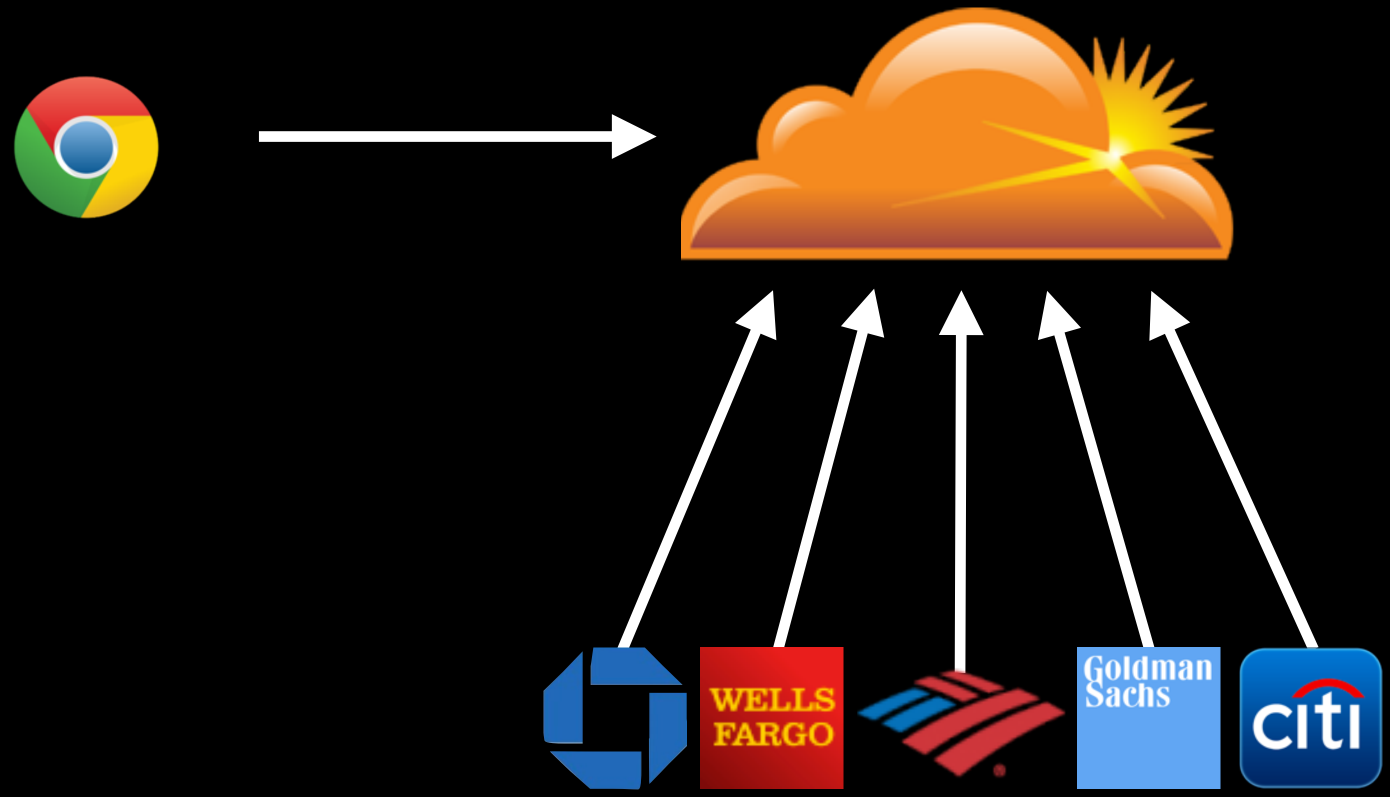
The PKI in today’s web



The PKI in today's web

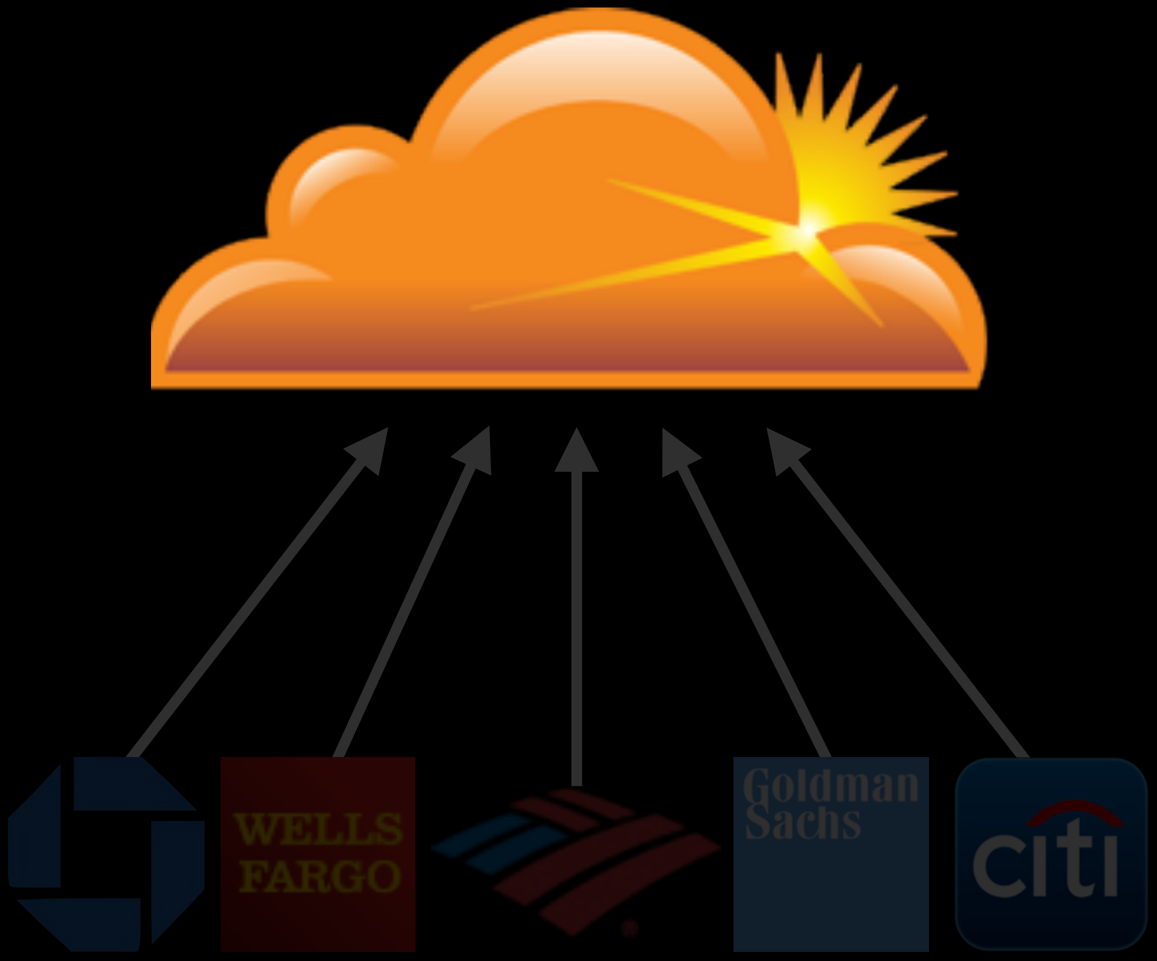


The PKI in today's web



Third-party Hosting Providers

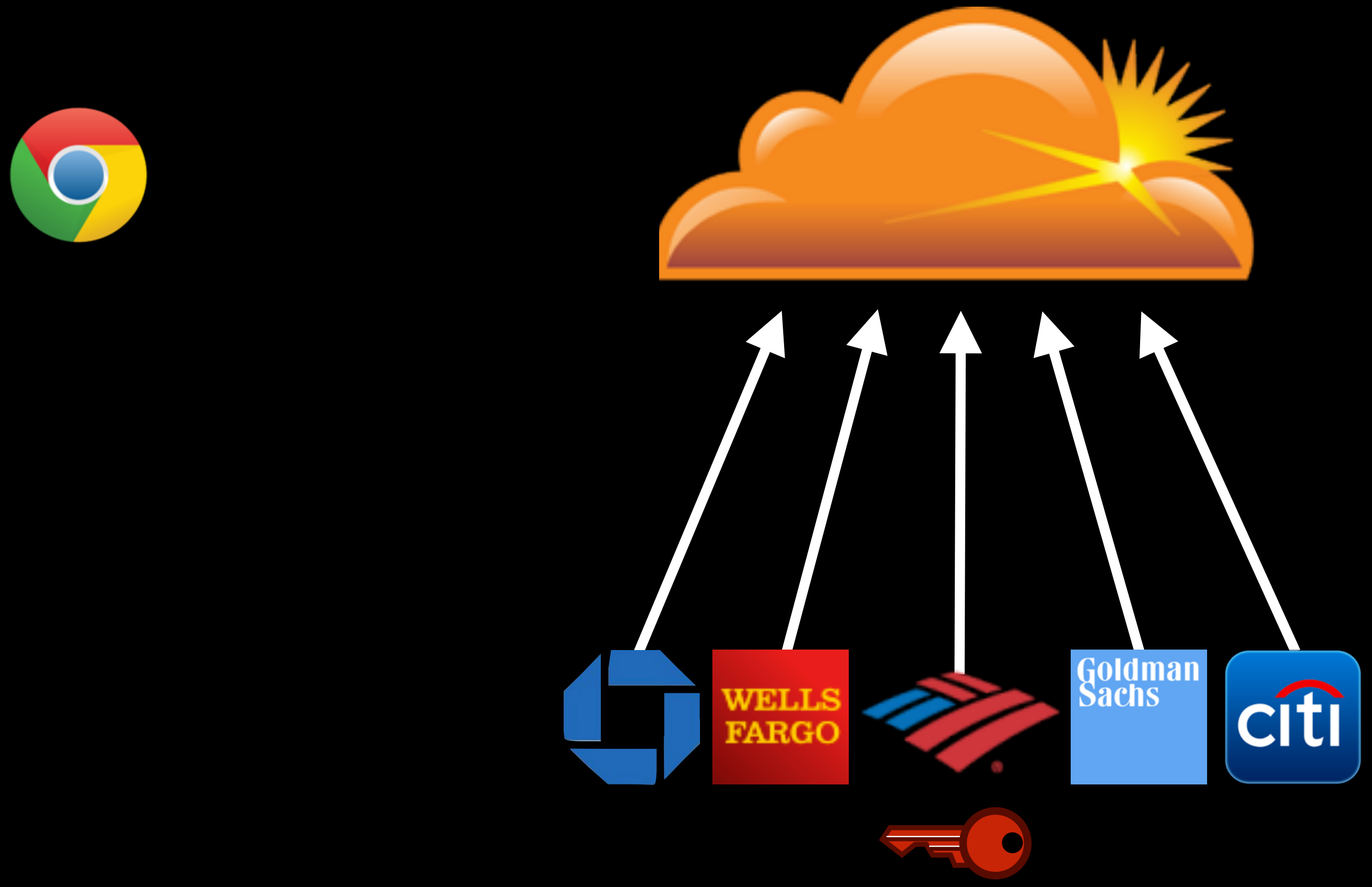
- Content delivery networks
- Web hosting services
- Cloud providers



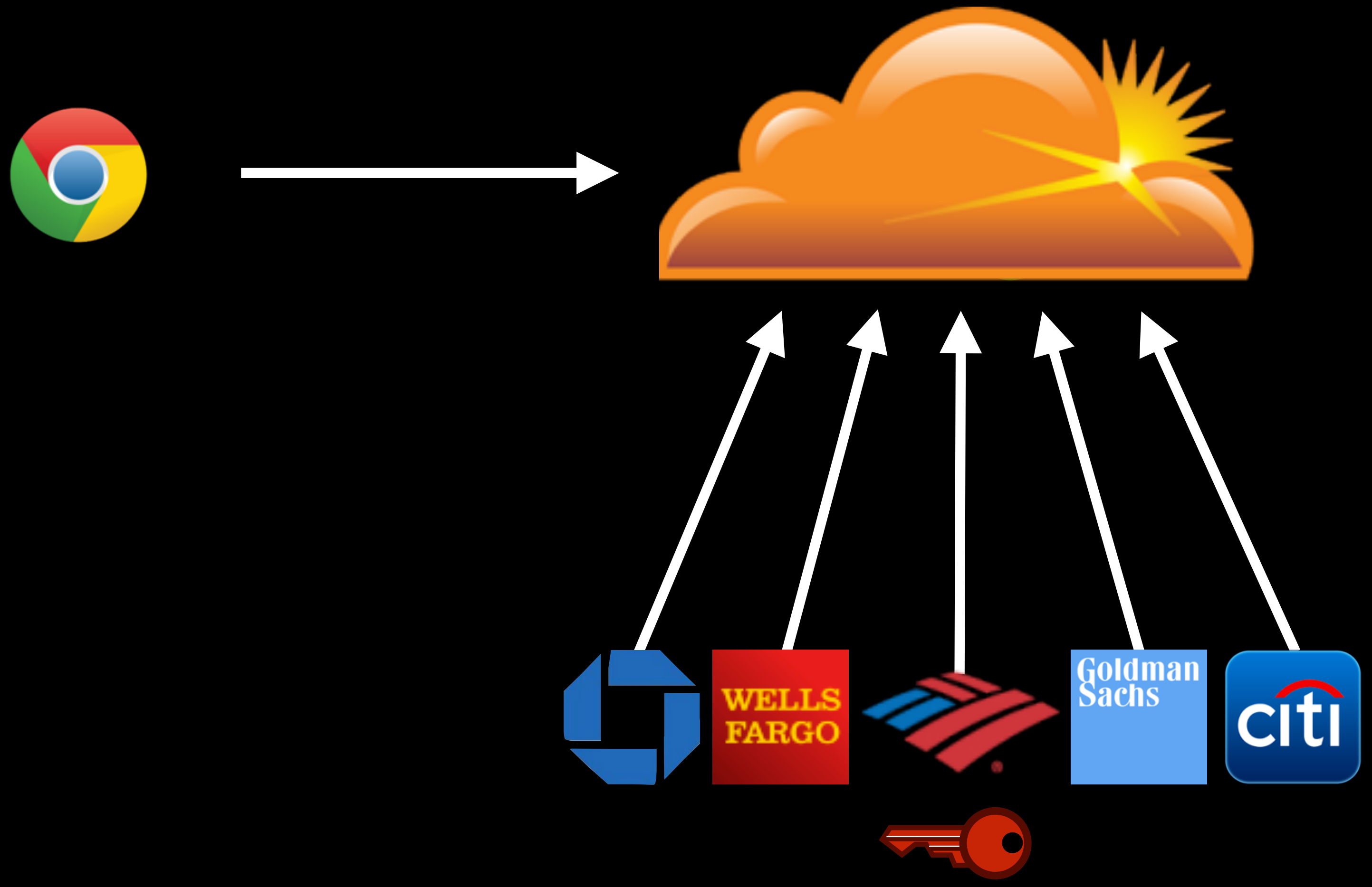
Varying levels of involvement

But all trusted to deliver content

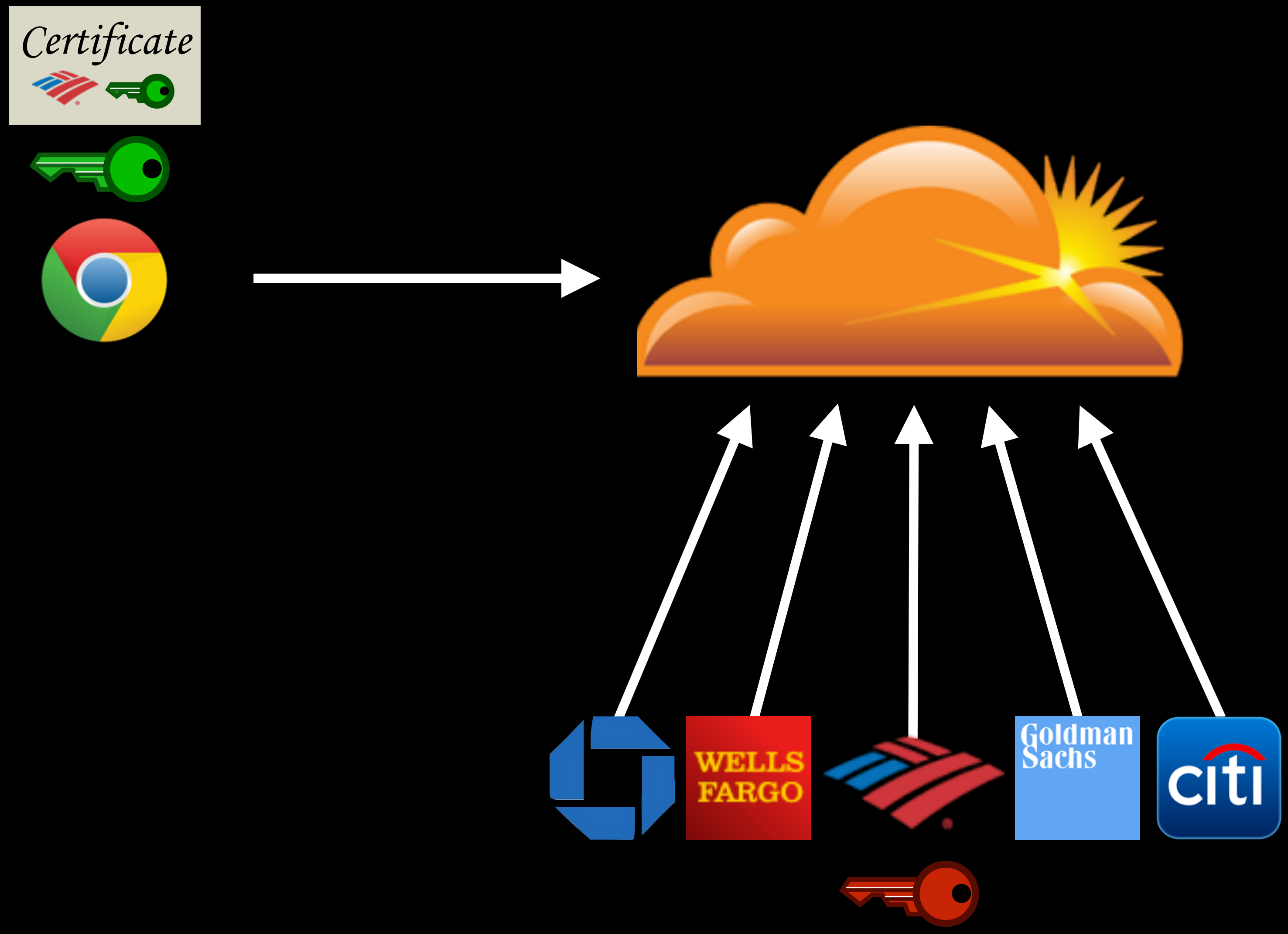
The PKI in today's web



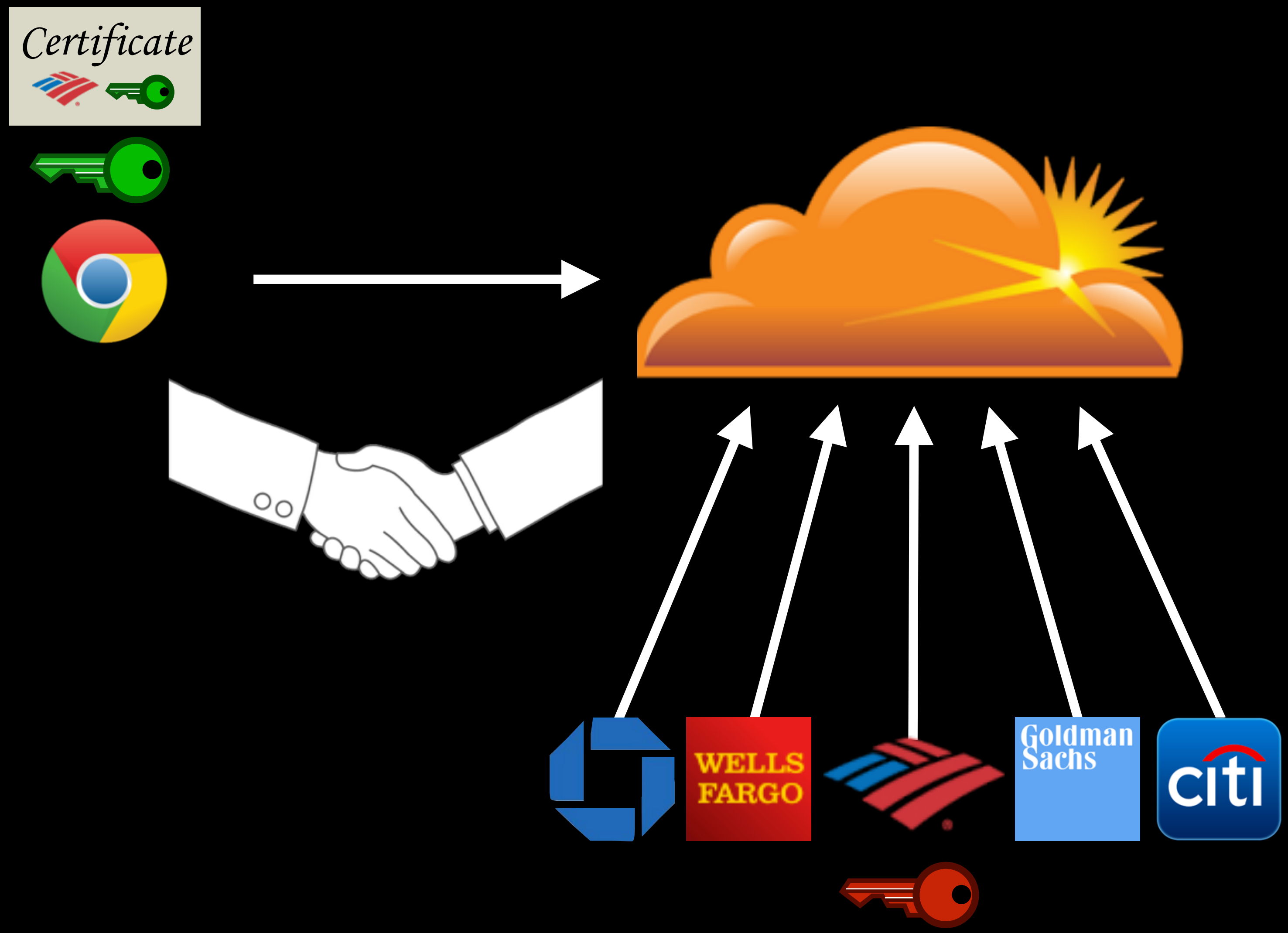
The PKI in today's web



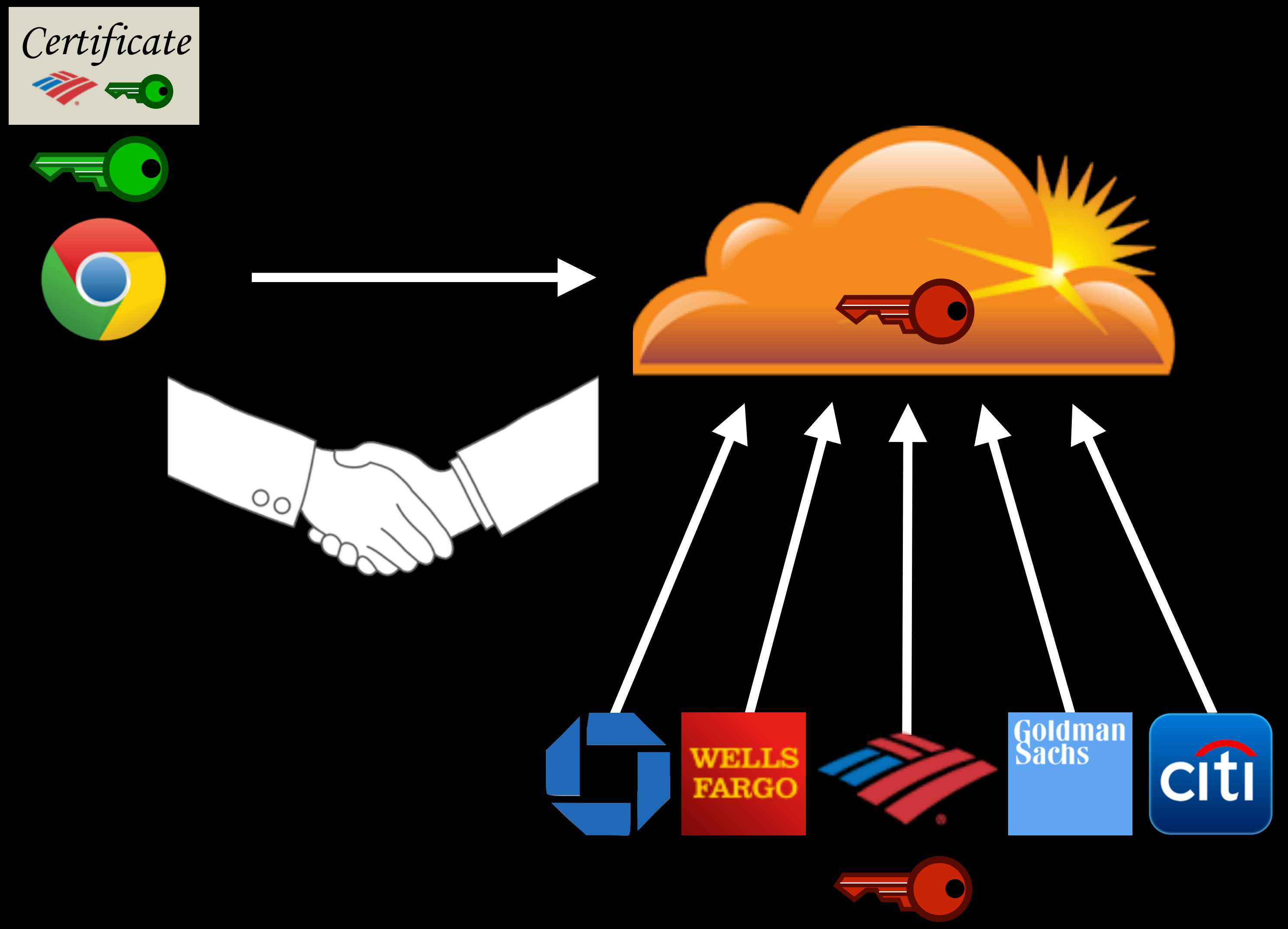
The PKI in today’s web



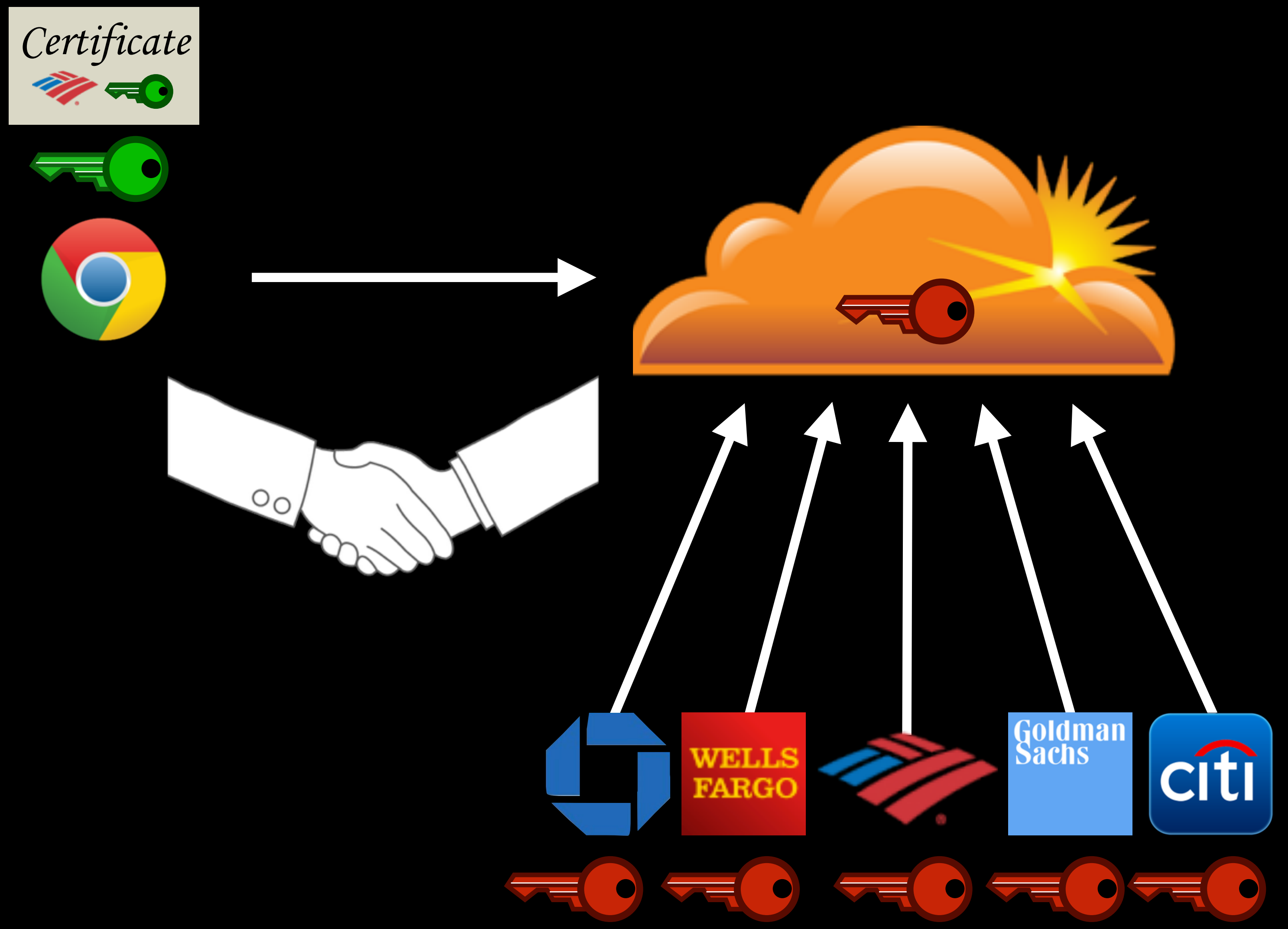
The PKI in today's web



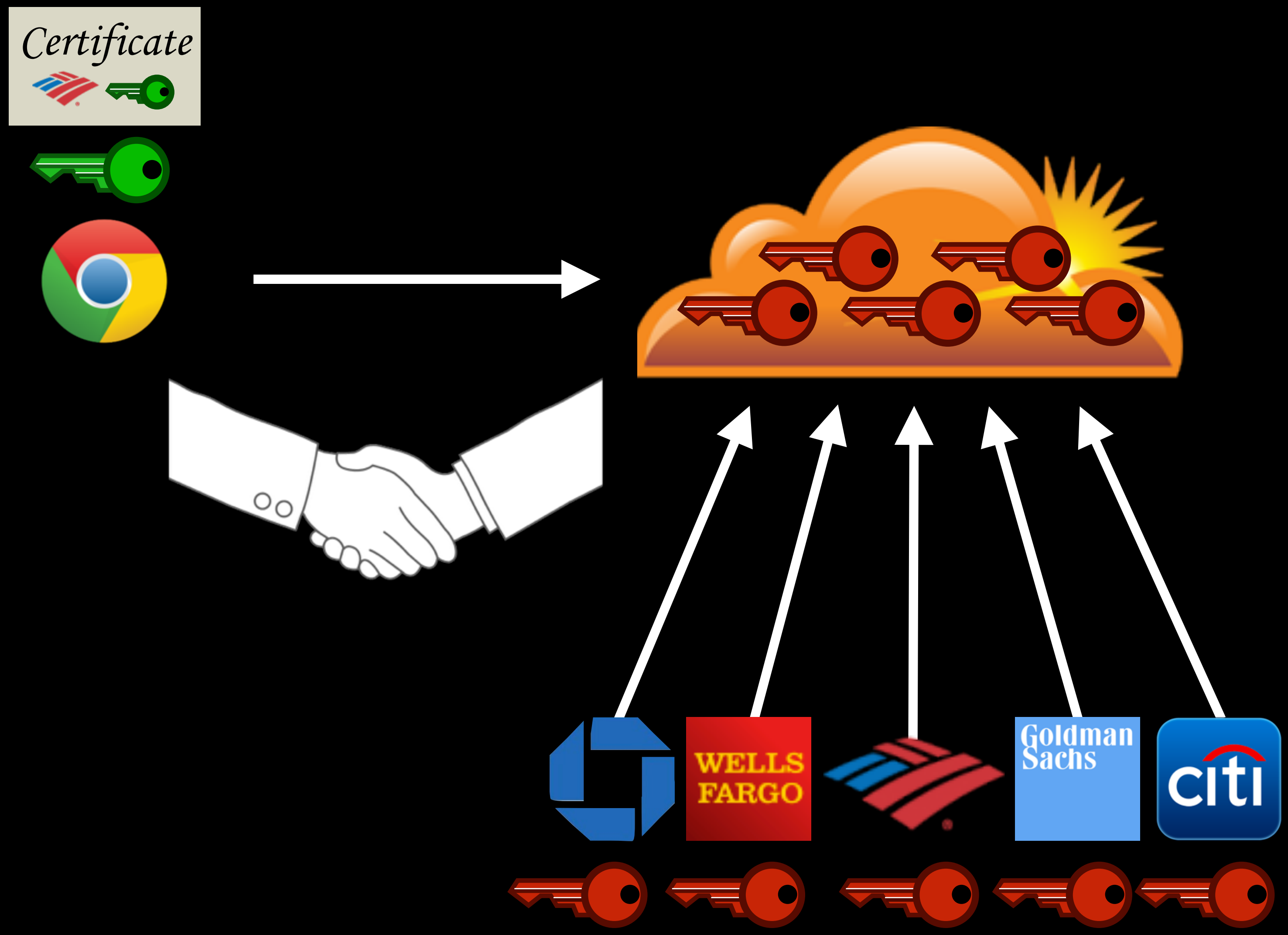
The PKI in today’s web



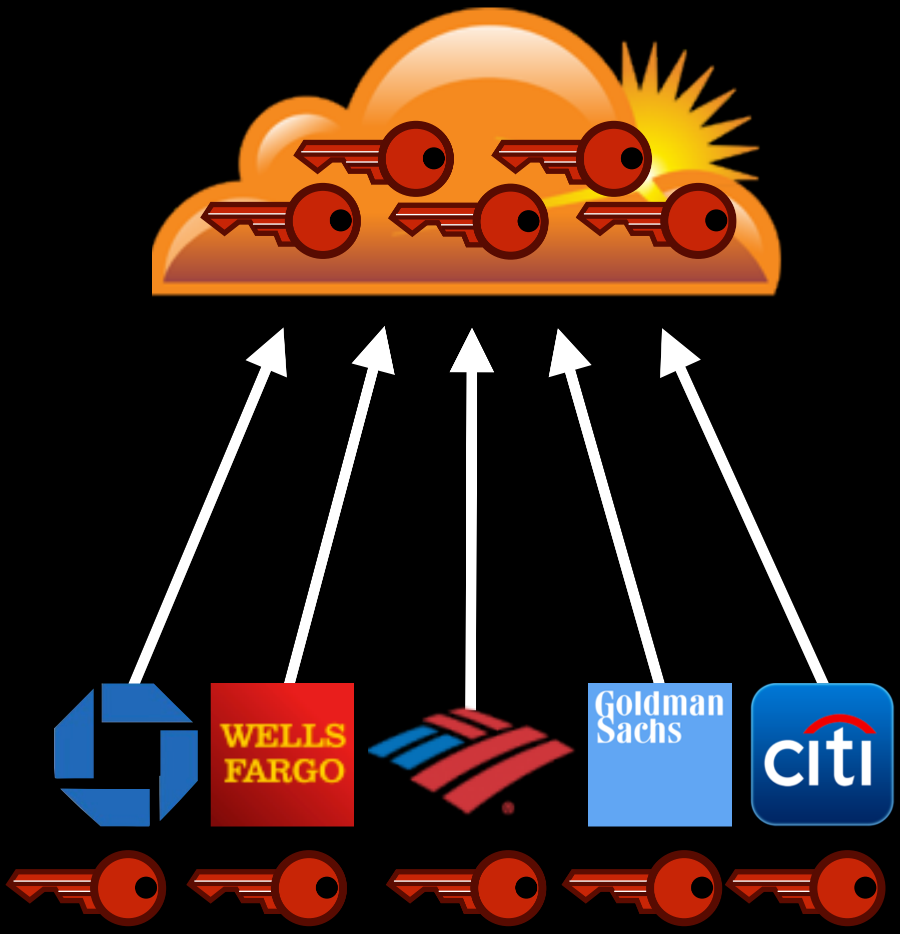
The PKI in today’s web



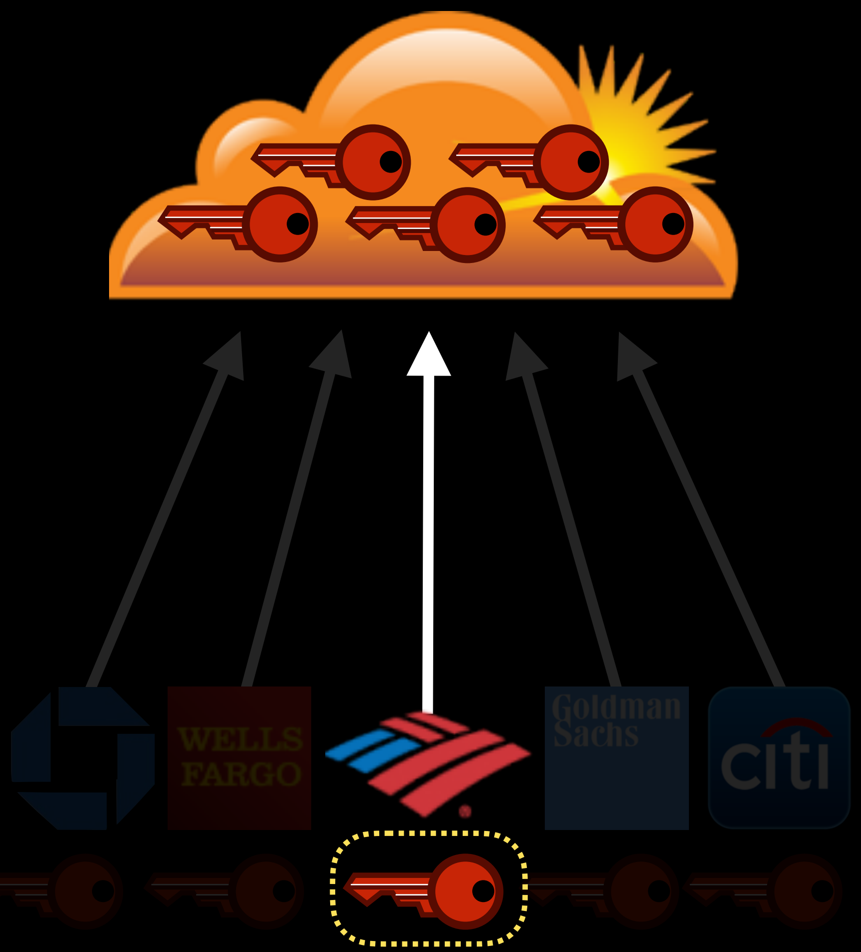
The PKI in today's web



Third-party hosting providers
know their customers’ private keys



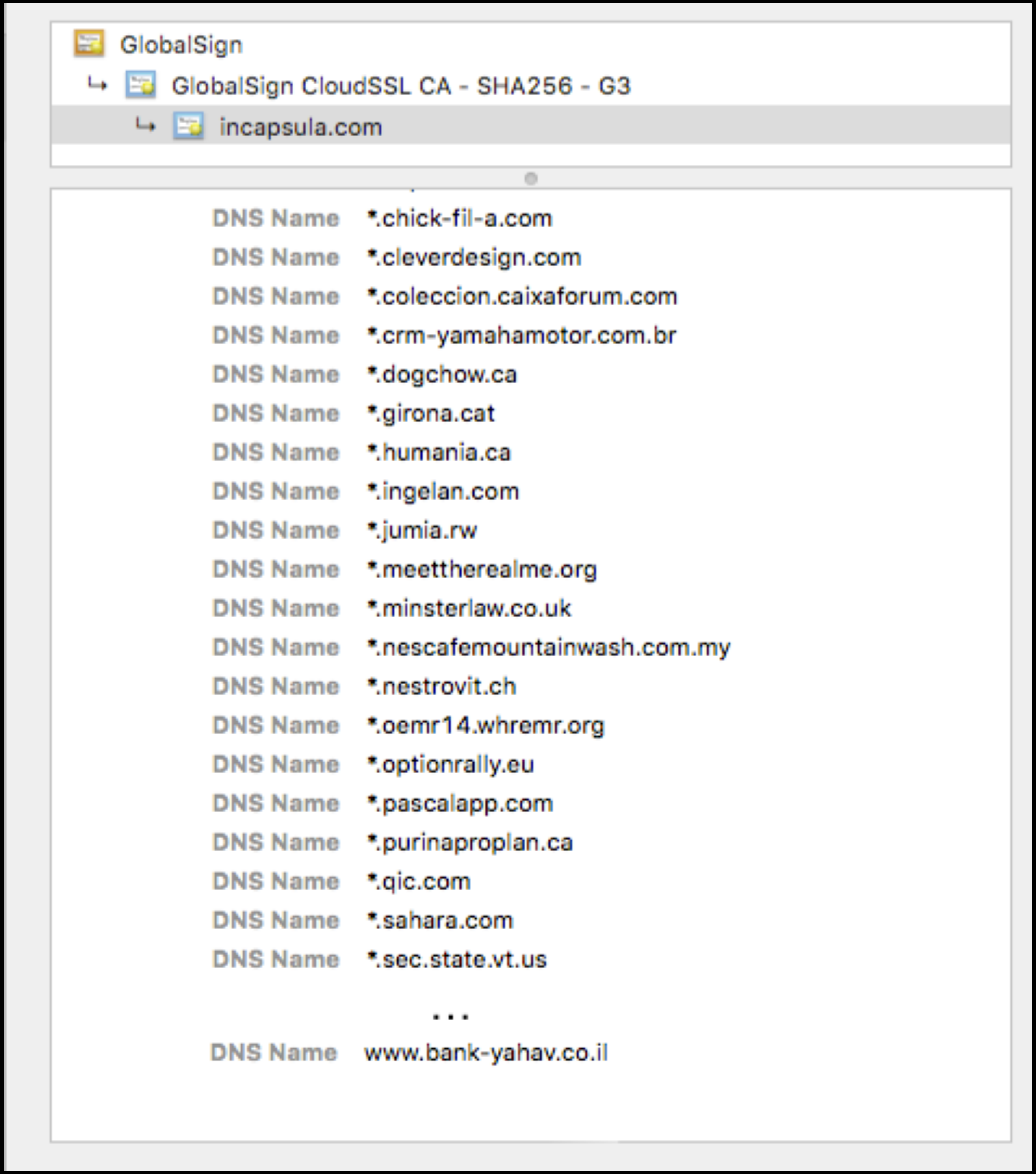
Third-party hosting providers **know their customers' private keys**



Authentication **fundamentally** assumes:

Only **knows** 🔑

Example of key sharing



GlobalSign
↳ GlobalSign CloudSSL CA - SHA256 - G3
↳ incapsula.com

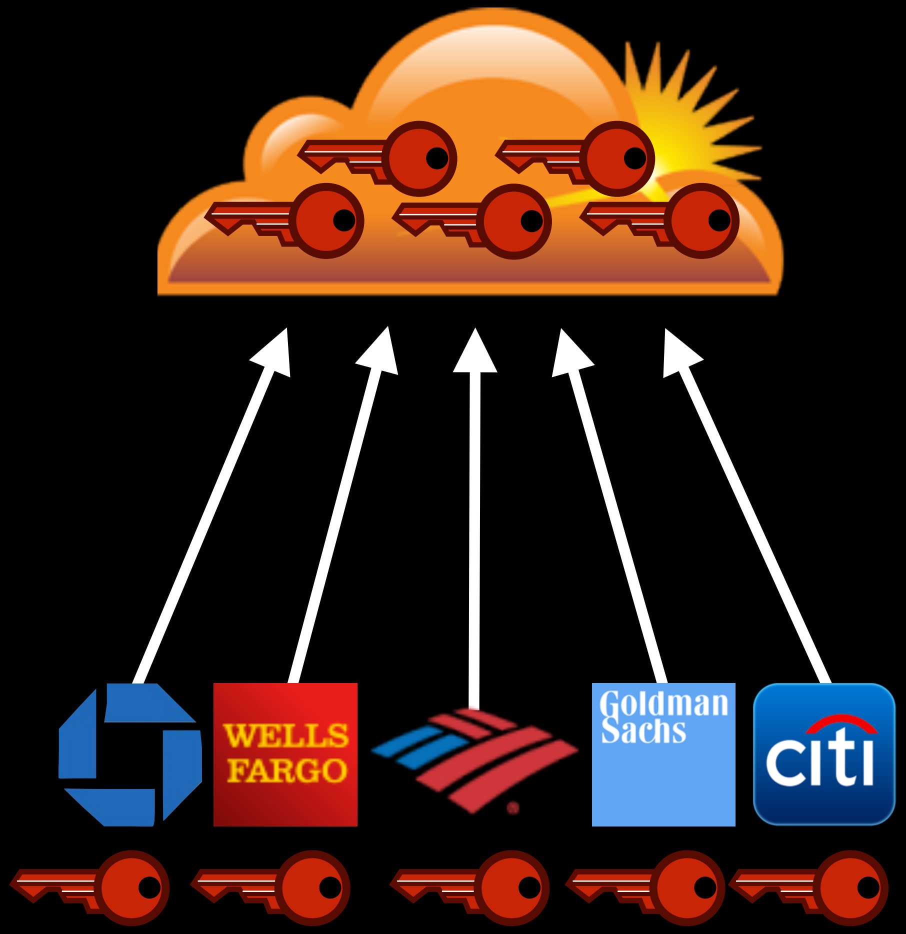
DNS Name	*.chick-fil-a.com
DNS Name	*.cleverdesign.com
DNS Name	*.coleccion.caixaforum.com
DNS Name	*.crm-yamahamotor.com.br
DNS Name	*.dogchow.ca
DNS Name	*.girona.cat
DNS Name	*.humania.ca
DNS Name	*.ingelan.com
DNS Name	*.jumia.rw
DNS Name	*.meettherealme.org
DNS Name	*.minsterlaw.co.uk
DNS Name	*.nescafemountainwash.com.my
DNS Name	*.nestrovit.ch
DNS Name	*.oemr14.whremr.org
DNS Name	*.optionrally.eu
DNS Name	*.pascalapp.com
DNS Name	*.purinaproplan.ca
DNS Name	*.qic.com
DNS Name	*.sahara.com
DNS Name	*.sec.state.vt.us
	. . .
DNS Name	www.bank-yahav.co.il

What’s wrong with sharing?

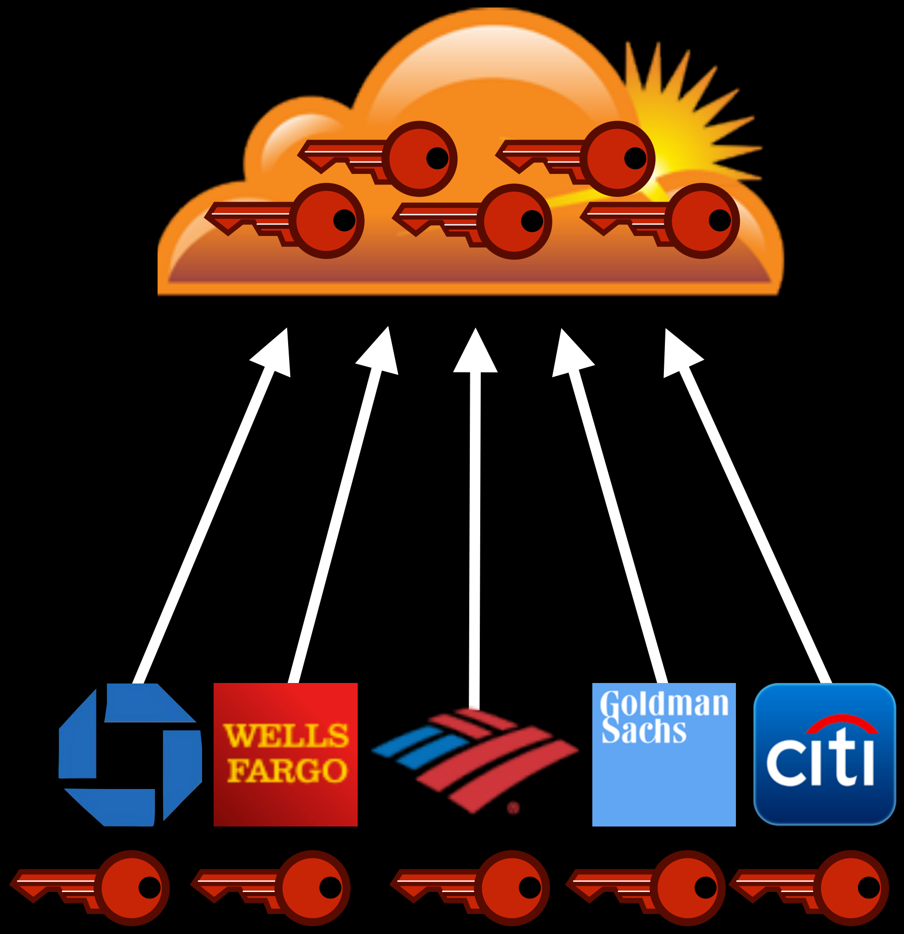
1. Complicates the trust model, users don’t know who they’re really trusting

2. Potential to create centralization of trust

3. Potential to create single point of failure (in terms of management)

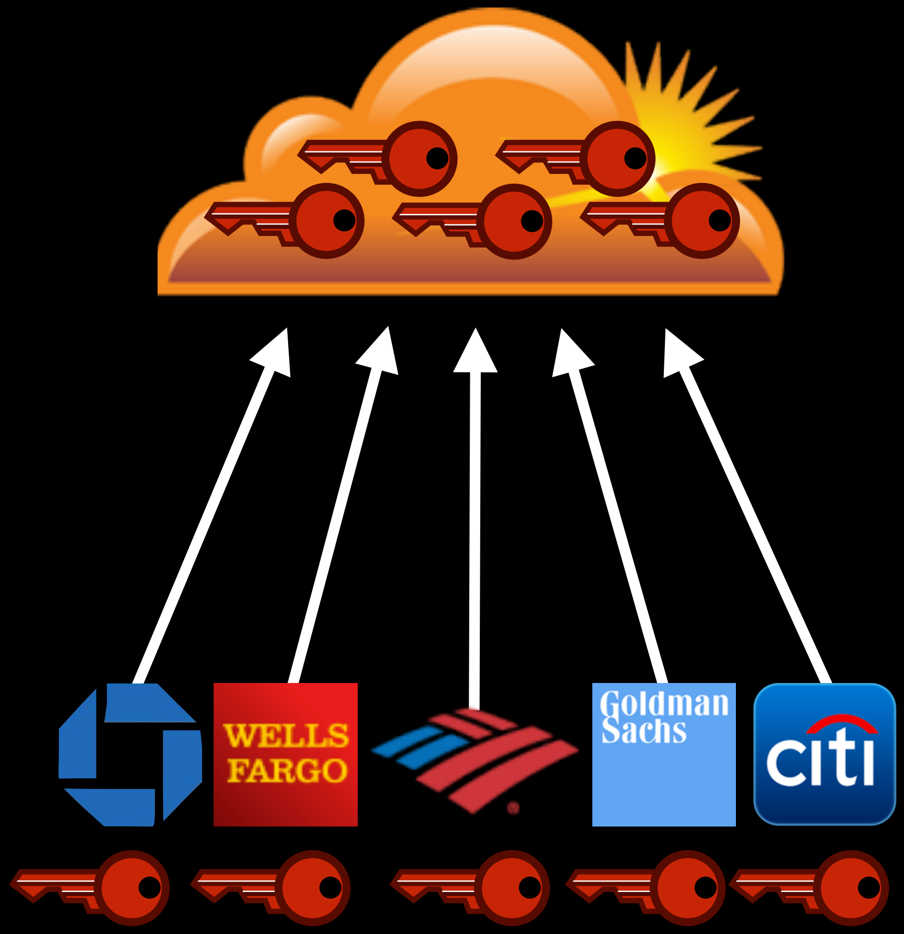


This study



This study

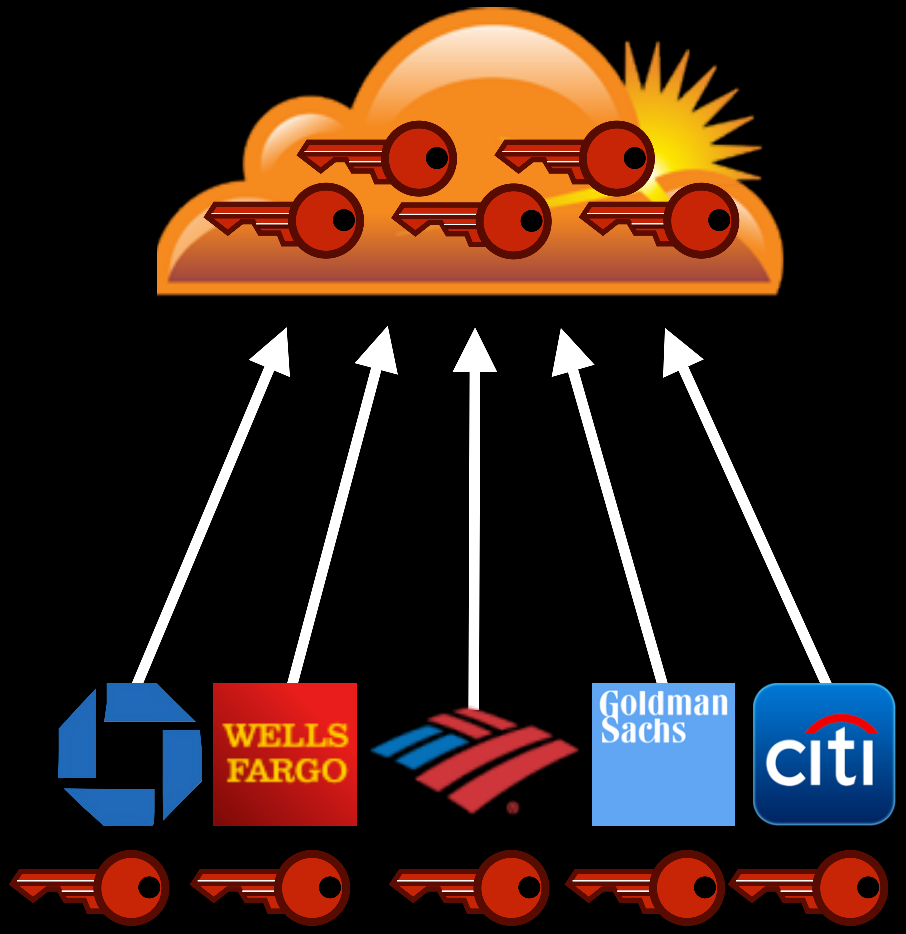
How many websites share their private keys?



This study

How many websites share their private keys?

How many keys have 3rd parties obtained?

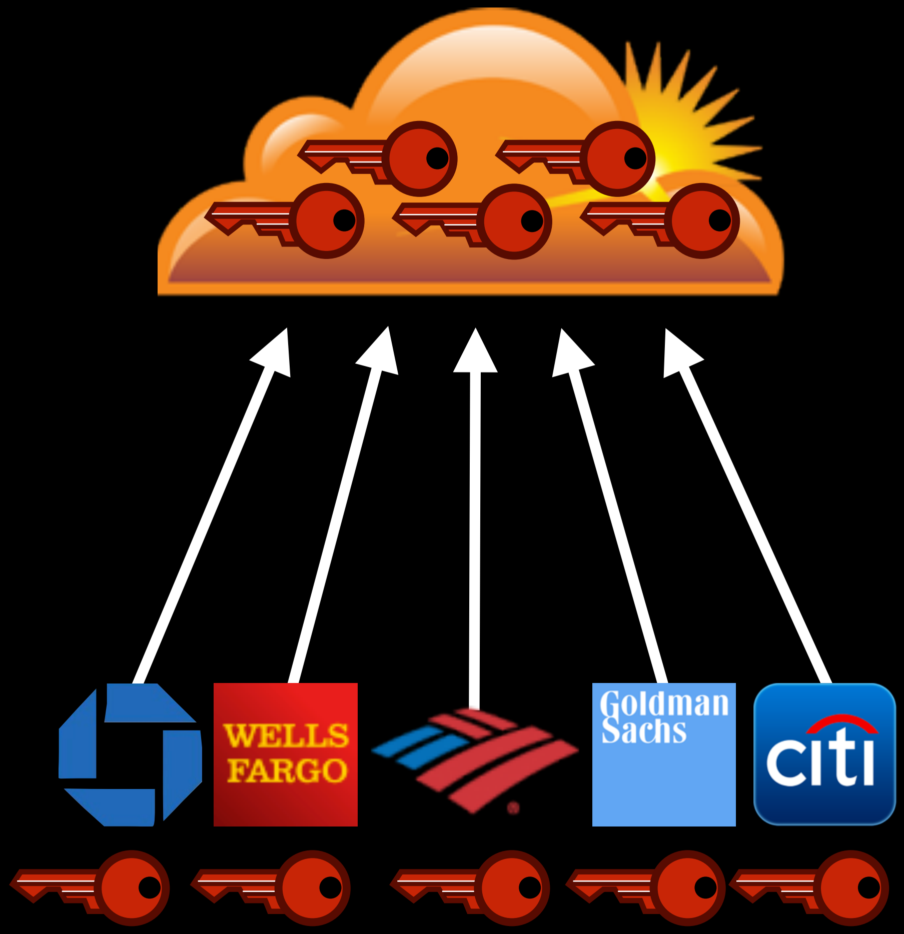


This study

How many websites share their private keys?

How many keys have 3rd parties obtained?

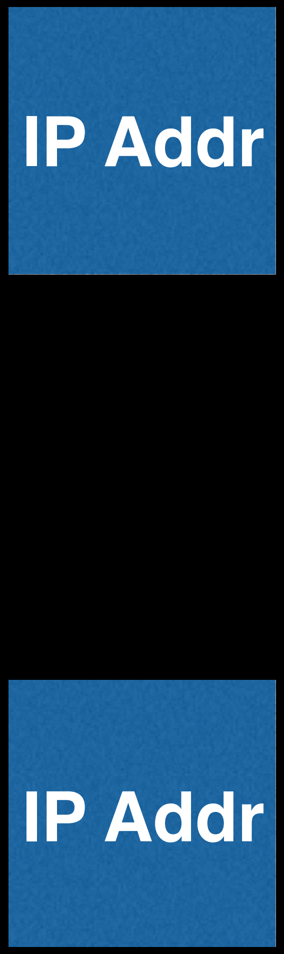
How has this affected key management?



How do we detect sharing *at scale*?

DATA

Rapid7 weekly port 443 scans 2013-2015



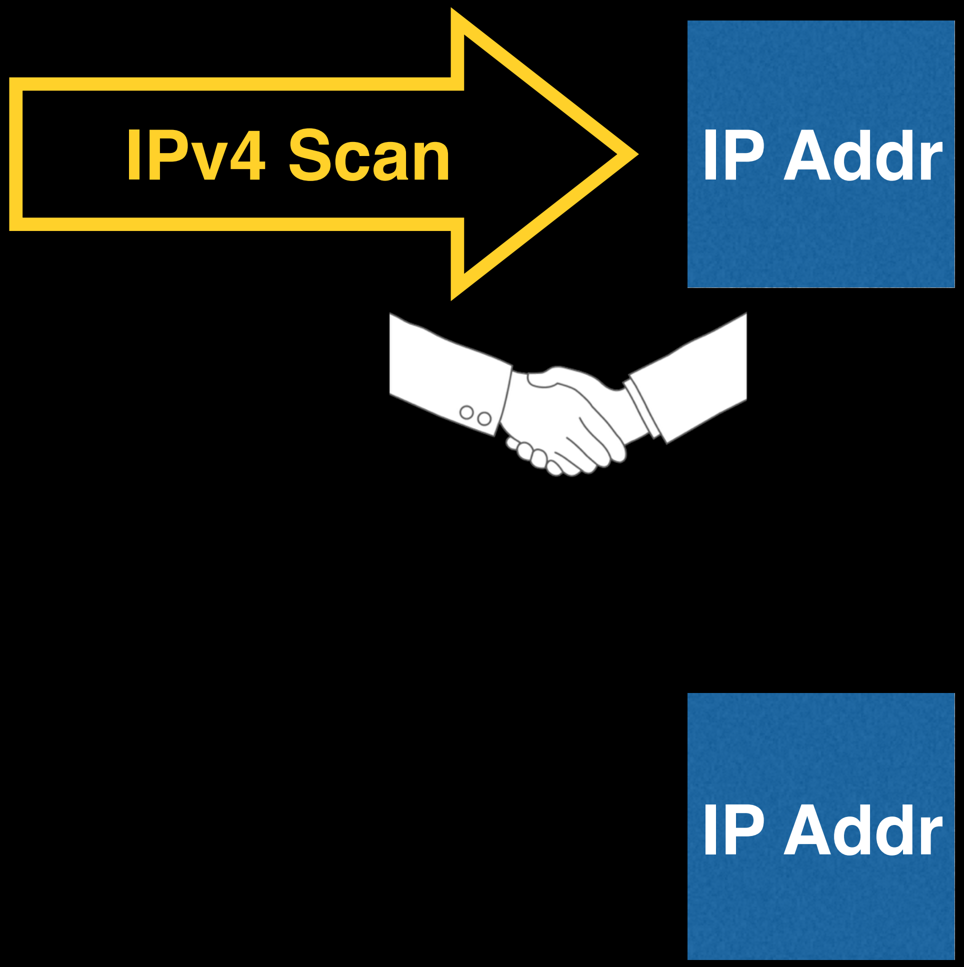
IP Addr

IP Addr

How do we detect sharing *at scale*?

DATA

Rapid7 weekly port 443 scans 2013-2015



IPv4 Scan

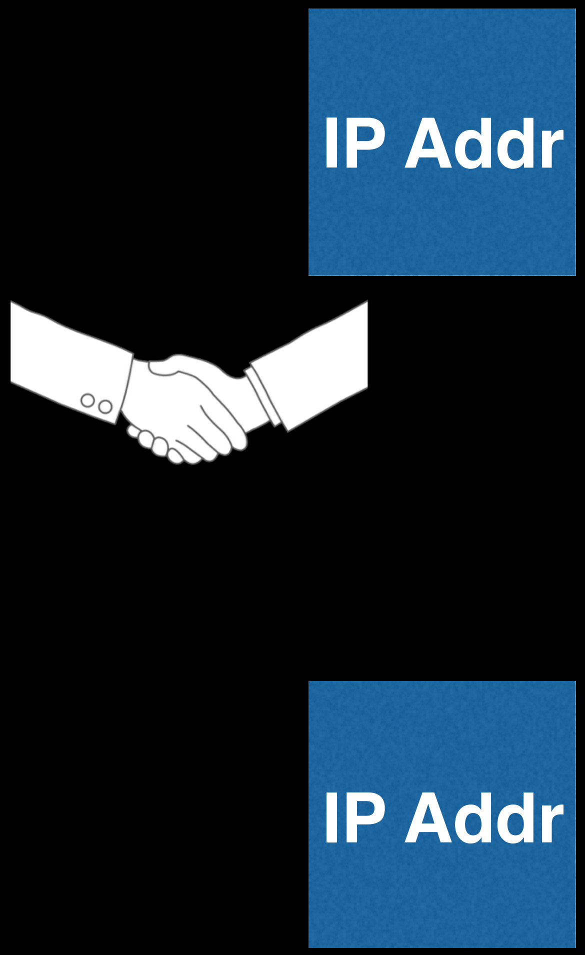
IP Addr

IP Addr

How do we detect sharing *at scale*?

DATA

Rapid7 weekly port 443 scans 2013-2015



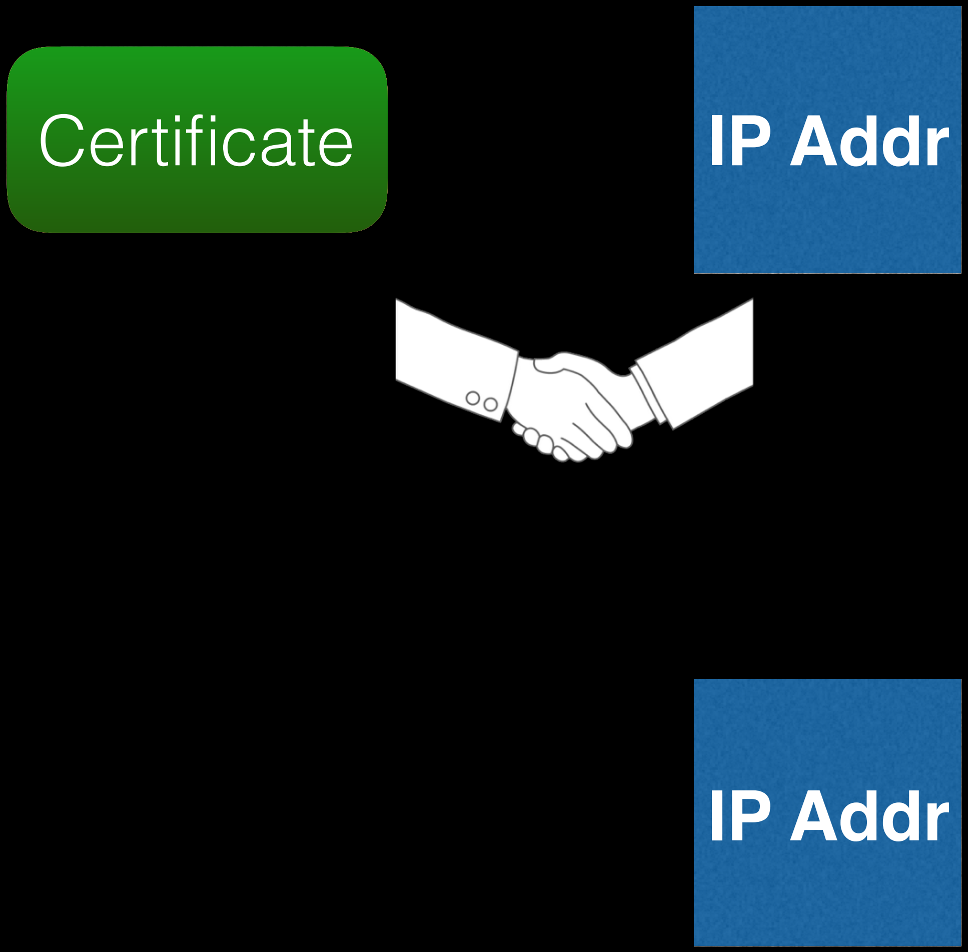
IP Addr

IP Addr

How do we detect sharing *at scale*?

DATA

Rapid7 weekly port 443 scans 2013-2015



Certificate

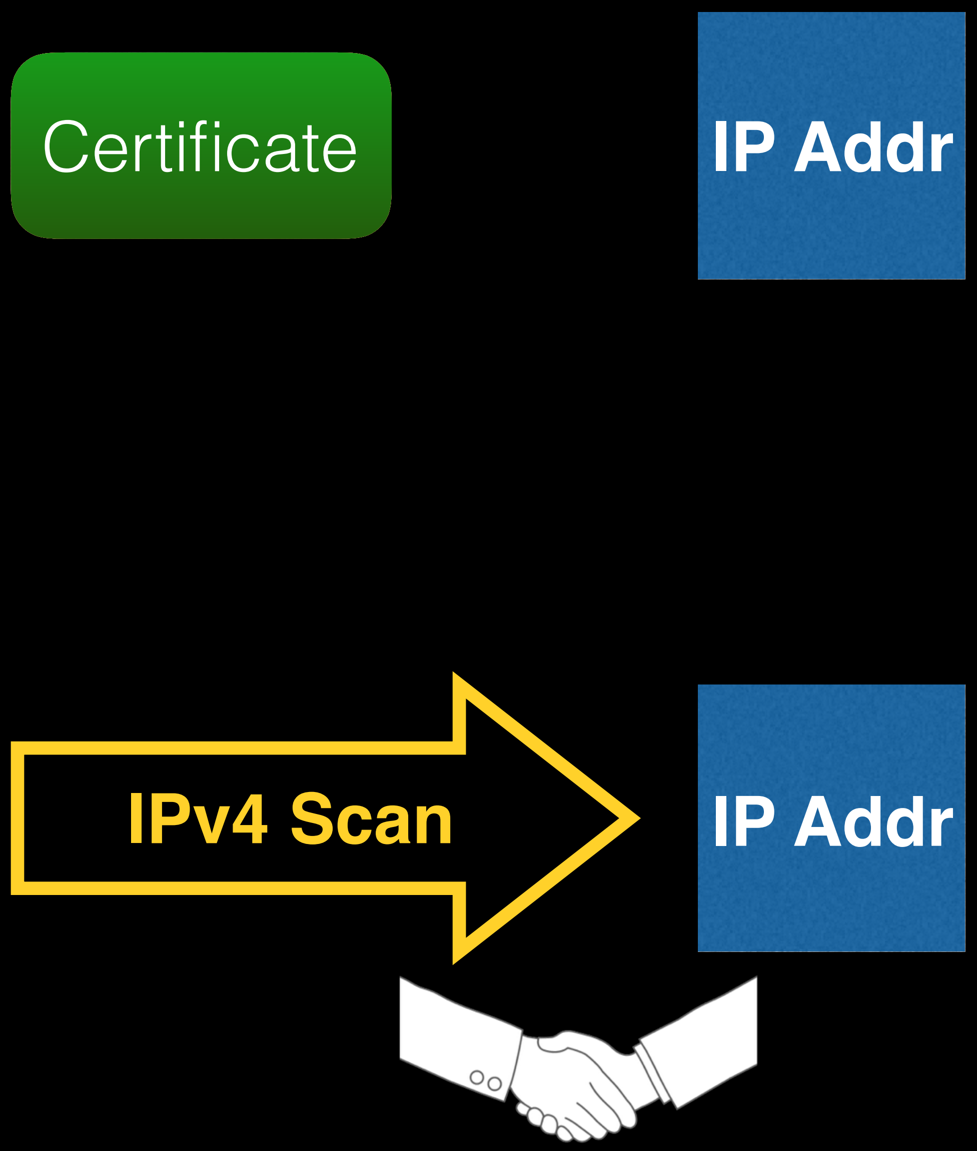
IP Addr

IP Addr

How do we detect sharing *at scale*?

DATA

Rapid7 weekly port 443 scans 2013-2015



Certificate

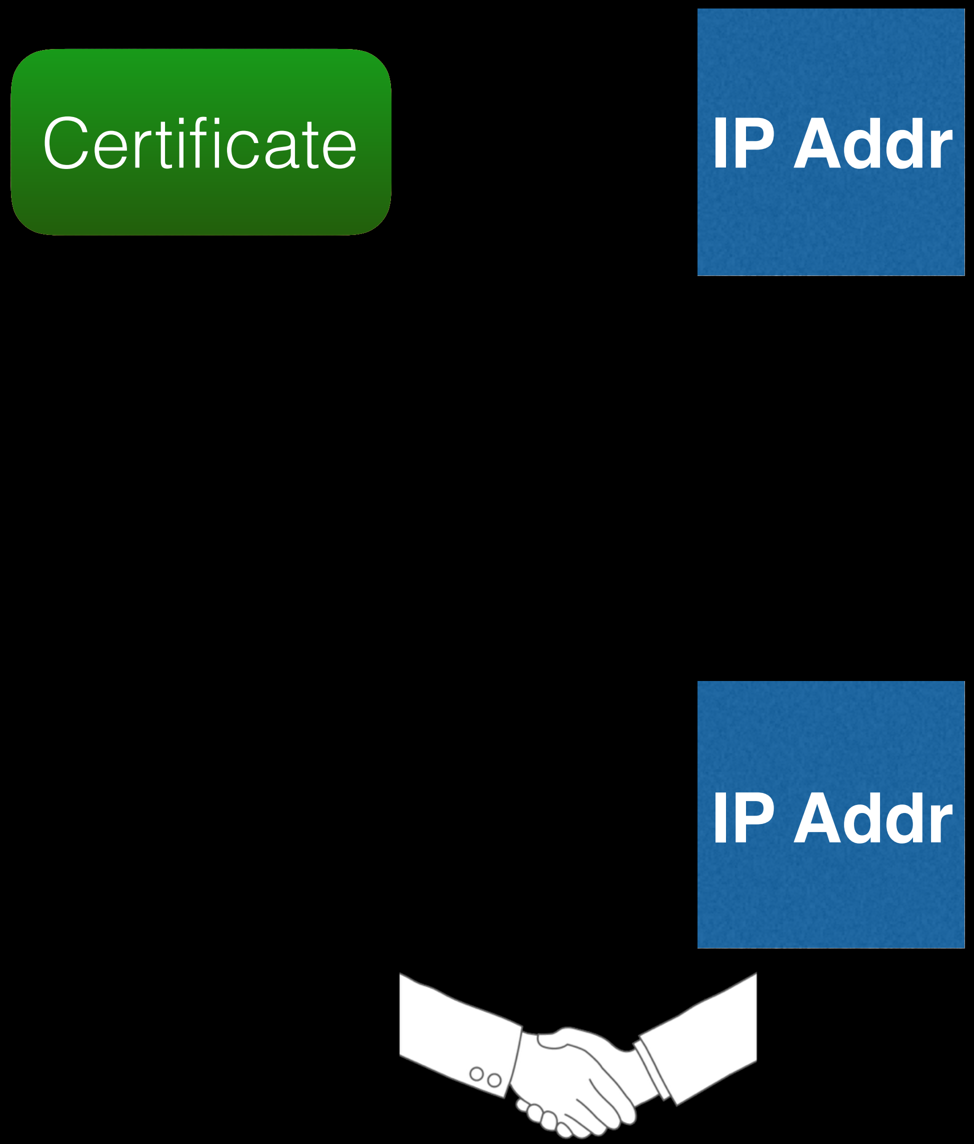
IP Addr

IPv4 Scan

IP Addr

How do we detect sharing *at scale*?

DATA Rapid7 weekly port 443 scans 2013-2015



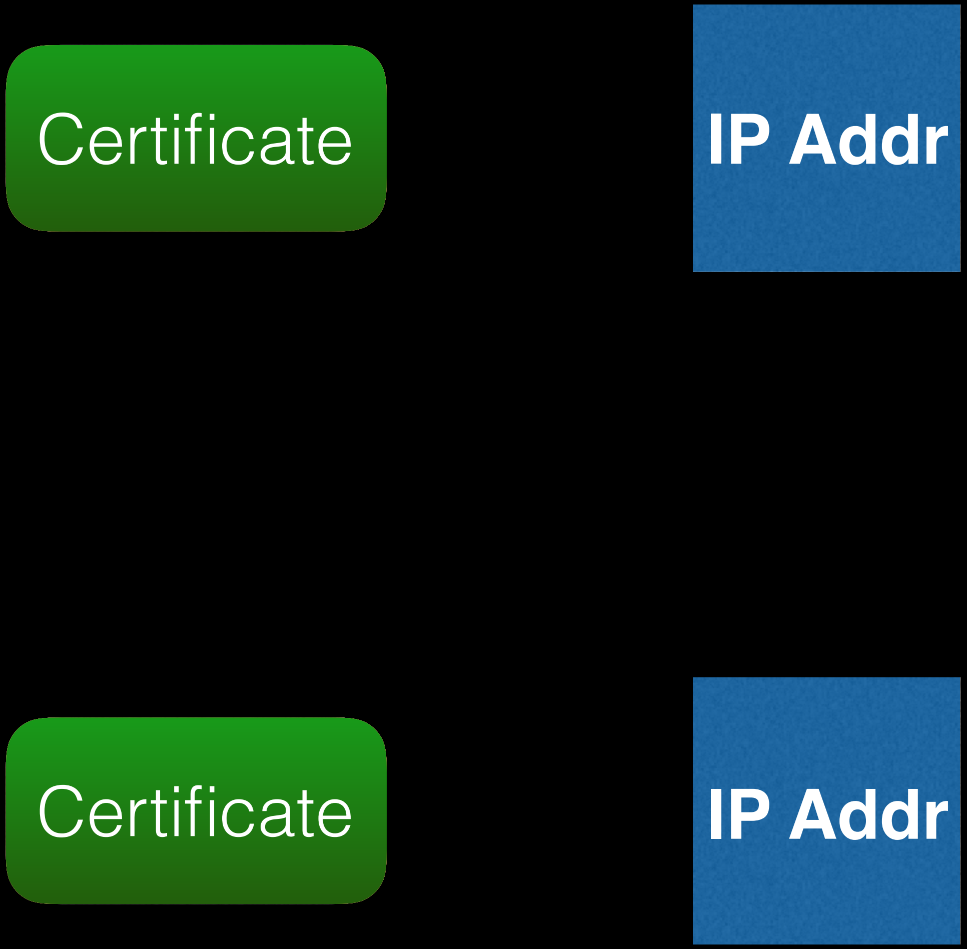
Certificate

IP Addr

IP Addr

How do we detect sharing *at scale*?

DATA Rapid7 weekly port 443 scans 2013-2015



Certificate

IP Addr

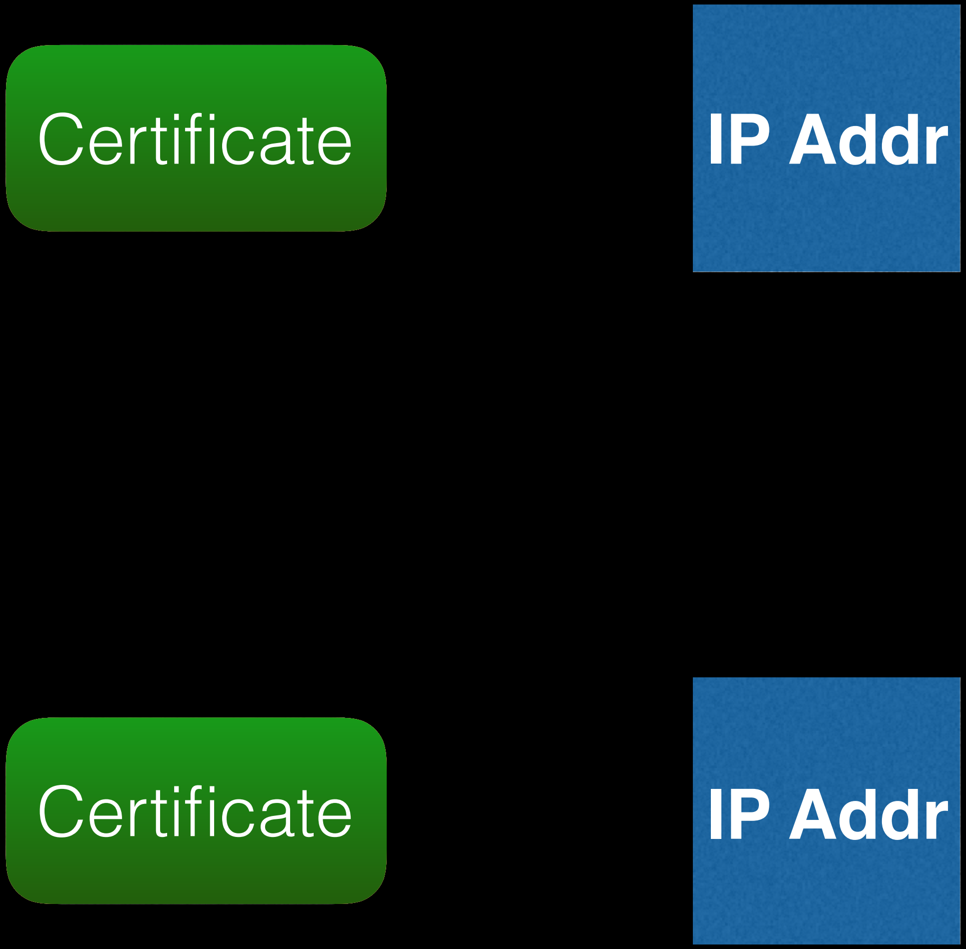
Certificate

IP Addr

How do we detect sharing *at scale*?

DATA

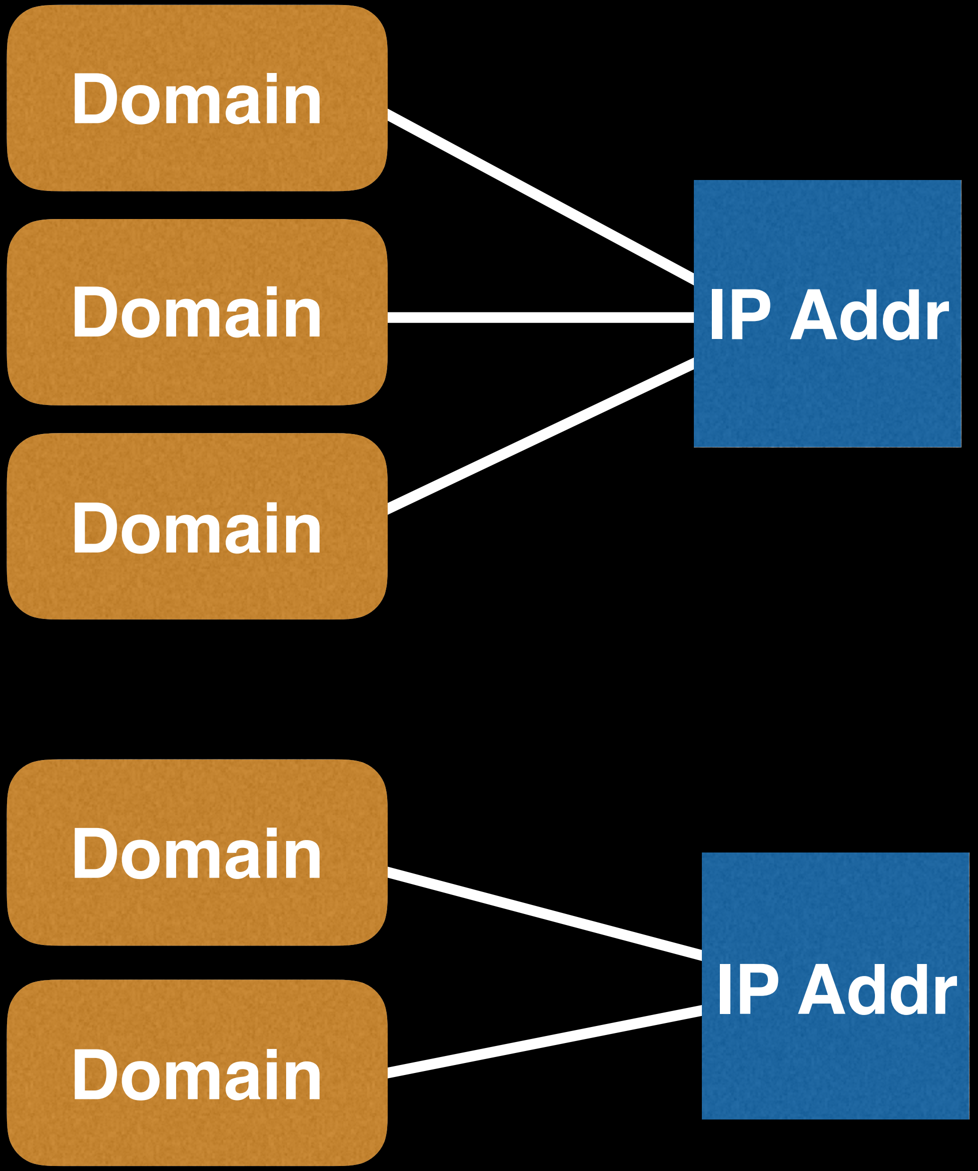
Rapid7 weekly port 443 scans 2013-2015
5.1 million valid leaf certificates



How do we detect sharing *at scale*?

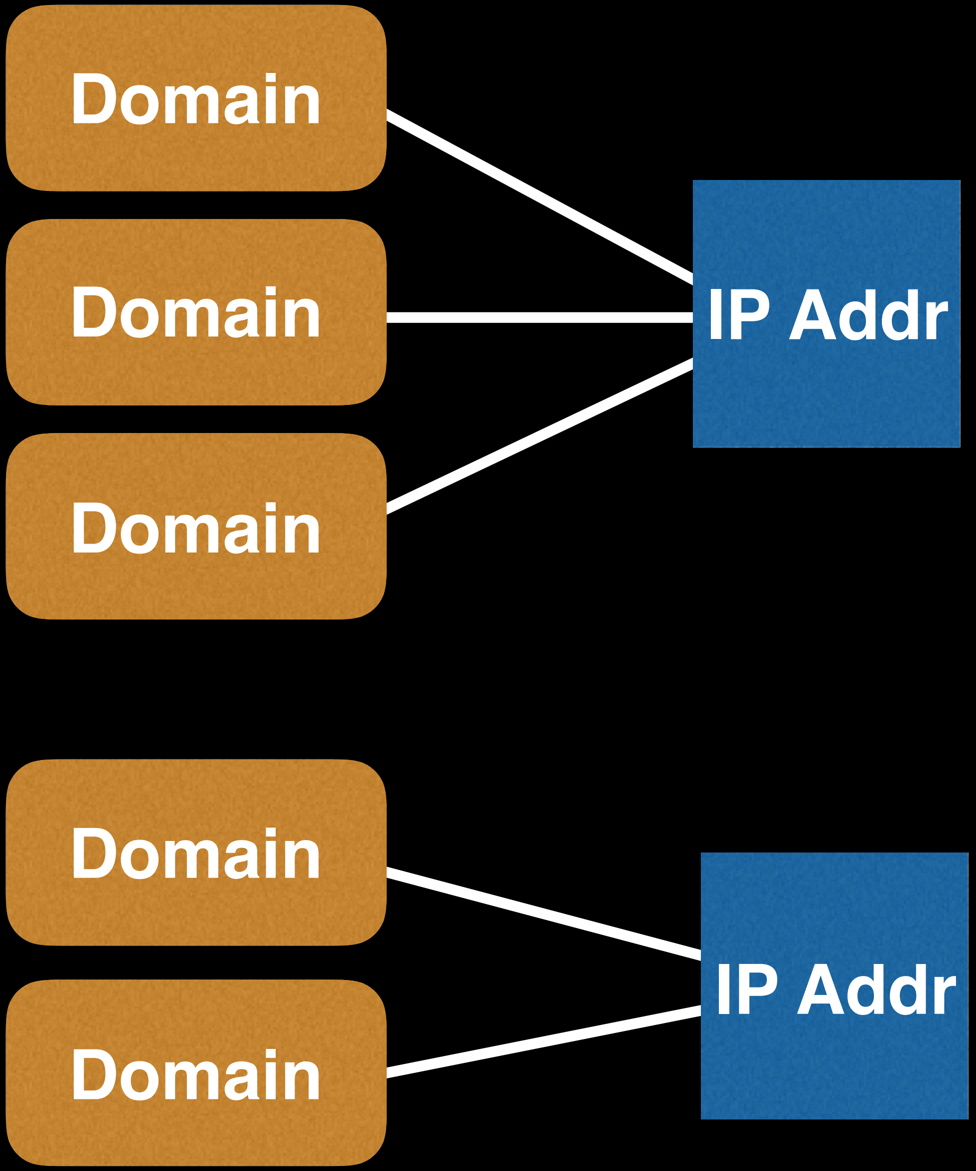
DATA

Rapid7 weekly port 443 scans 2013-2015
5.1 million valid leaf certificates



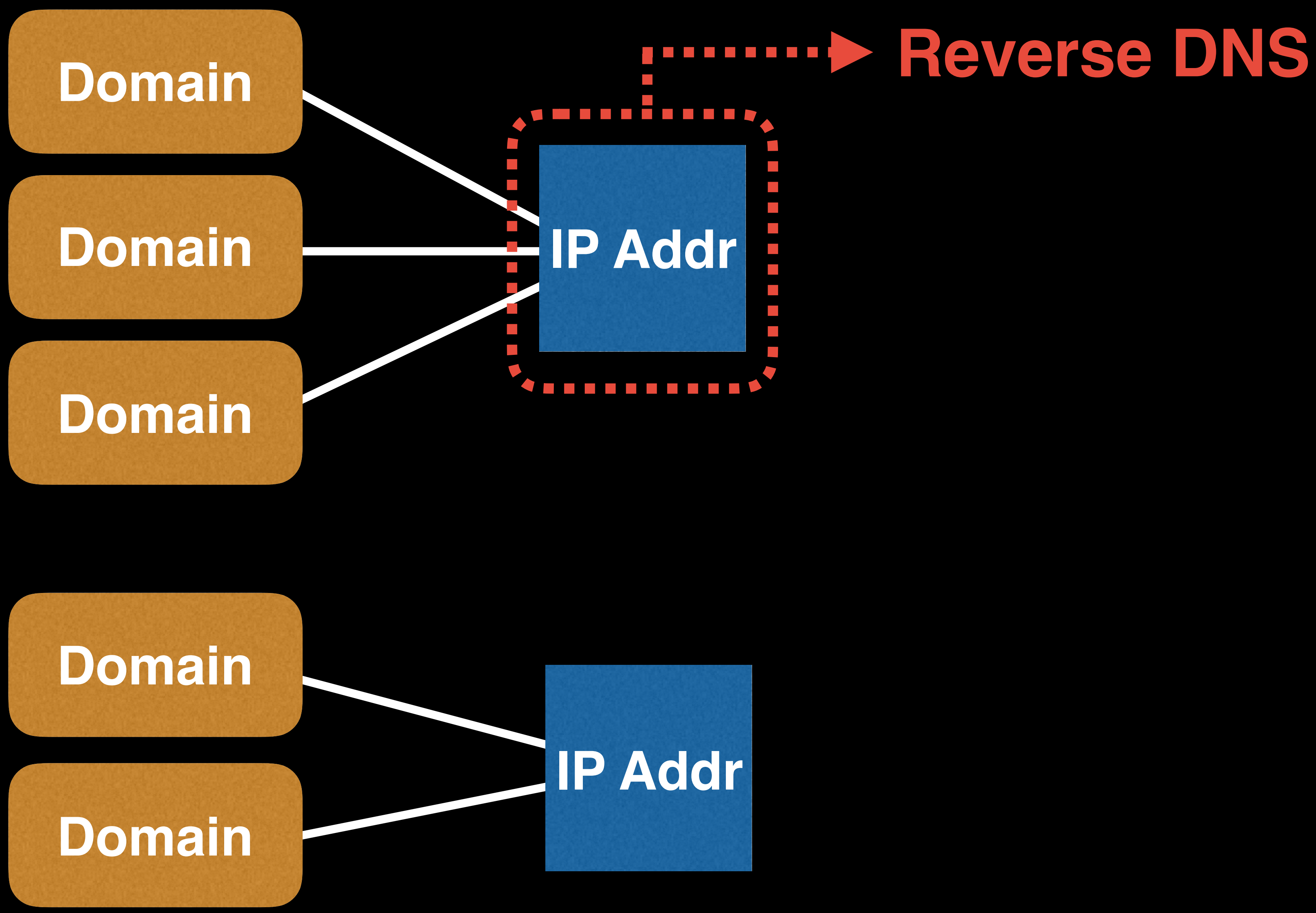
How do we detect sharing *at scale*?

Does the same entity that owns the domain own and operate the server at that IP address?



How do we detect sharing *at scale*?

Does the same entity that owns the domain own and operate the server at that IP address?



Domain equivalence?

google.com	google.co.uk	
google.com	youtube.com	
nestle.com	friskies.com	
whitehouse.gov	whitehouse.com	

Domain names alone are *not enough*

Incorporating whois

Emails in whois records reflect administrative domain

Incorporating whois

```
google.com

google.co.uk

google.de

zagat.com

golang.org
```

Emails in whois records reflect administrative domain

Incorporating whois

```
whois  google.com

       google.co.uk

       google.de

       zagat.com

       golang.org
```

```
Registrant Email: dns-admin@google.com

Admin Email: dns-admin@google.com

Tech Email: dns-admin@google.com
```

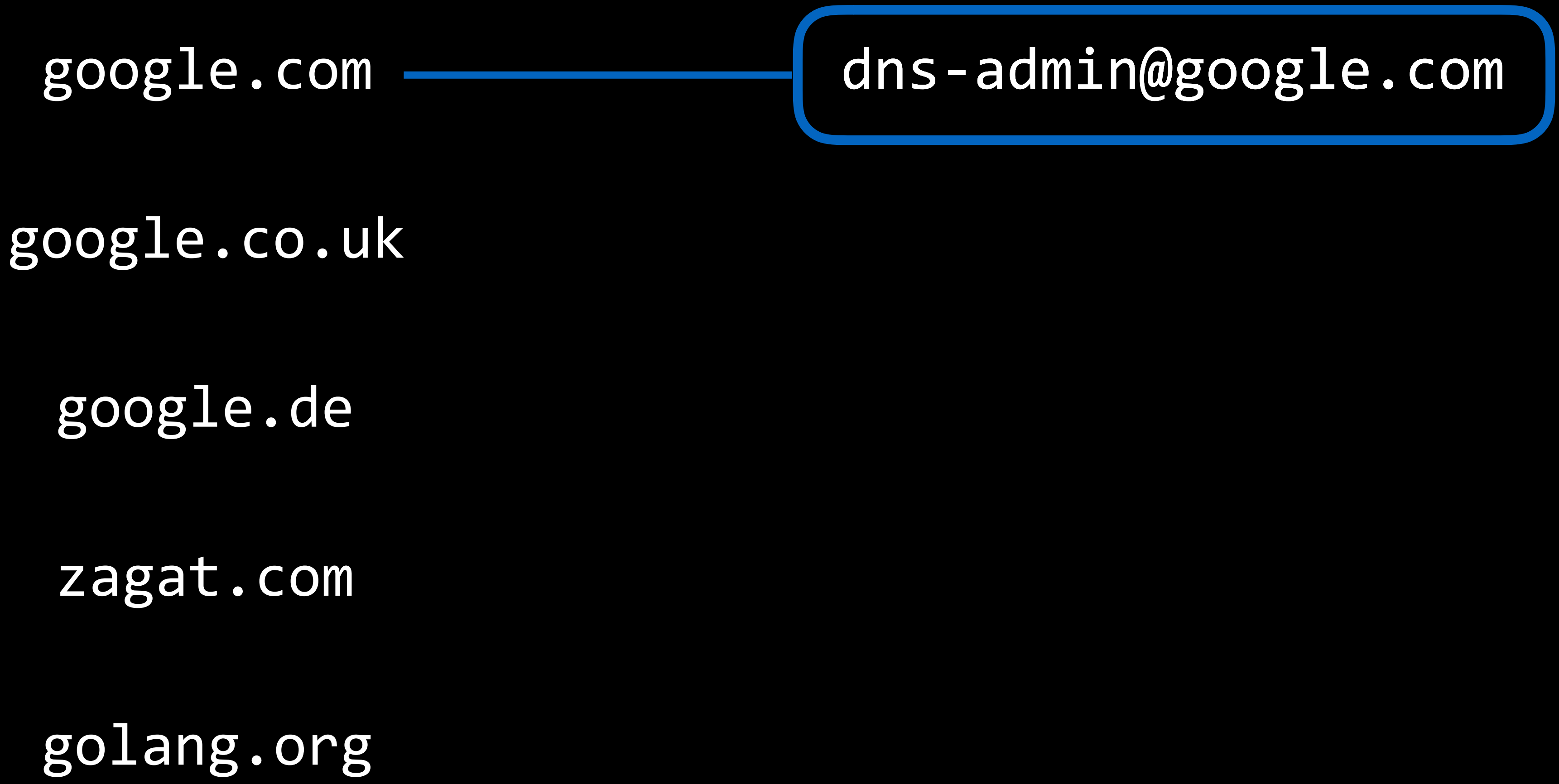
Emails in whois records reflect administrative domain

Incorporating whois

google.com	dns-admin@google.com
google.co.uk	
google.de	
zagat.com	
golang.org	

Emails in whois records reflect administrative domain

Incorporating whois



google.com

dns-admin@google.com

google.co.uk

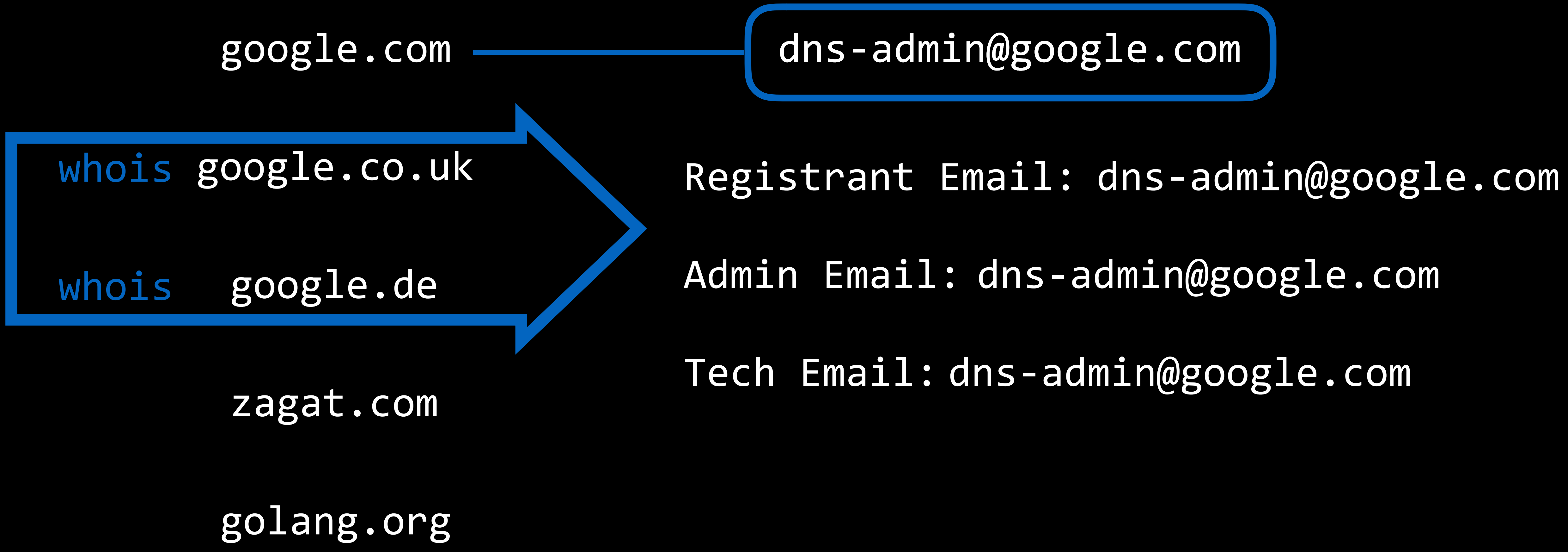
google.de

zagat.com

golang.org

Emails in whois records reflect administrative domain

Incorporating whois



google.com

dns-admin@google.com

```
whois google.co.uk

whois   google.de
```

zagat.com

golang.org

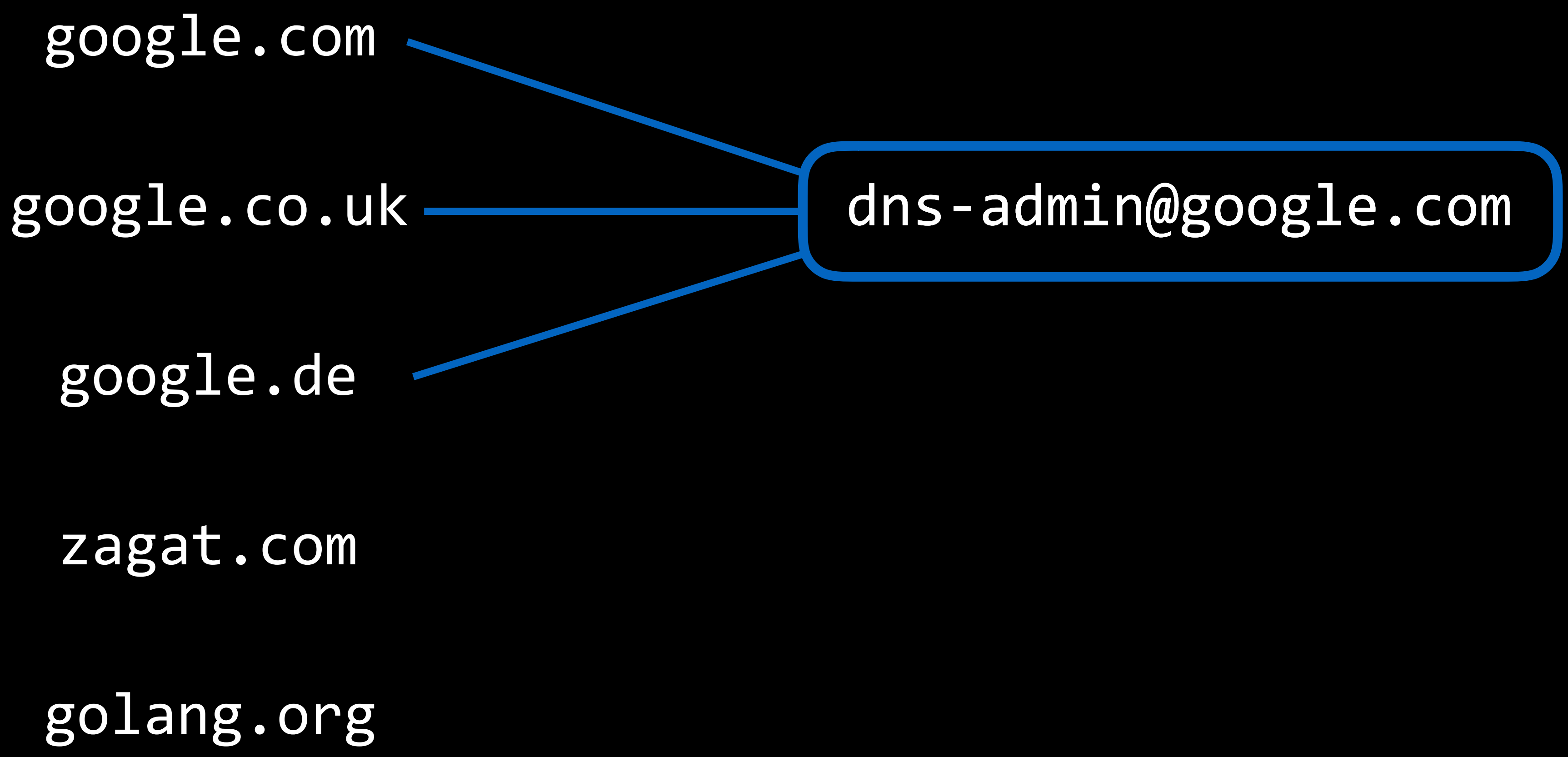
```
Registrant Email: dns-admin@google.com

Admin Email: dns-admin@google.com

Tech Email: dns-admin@google.com
```

Emails in whois records reflect administrative domain

Incorporating whois



google.com

google.co.uk

google.de

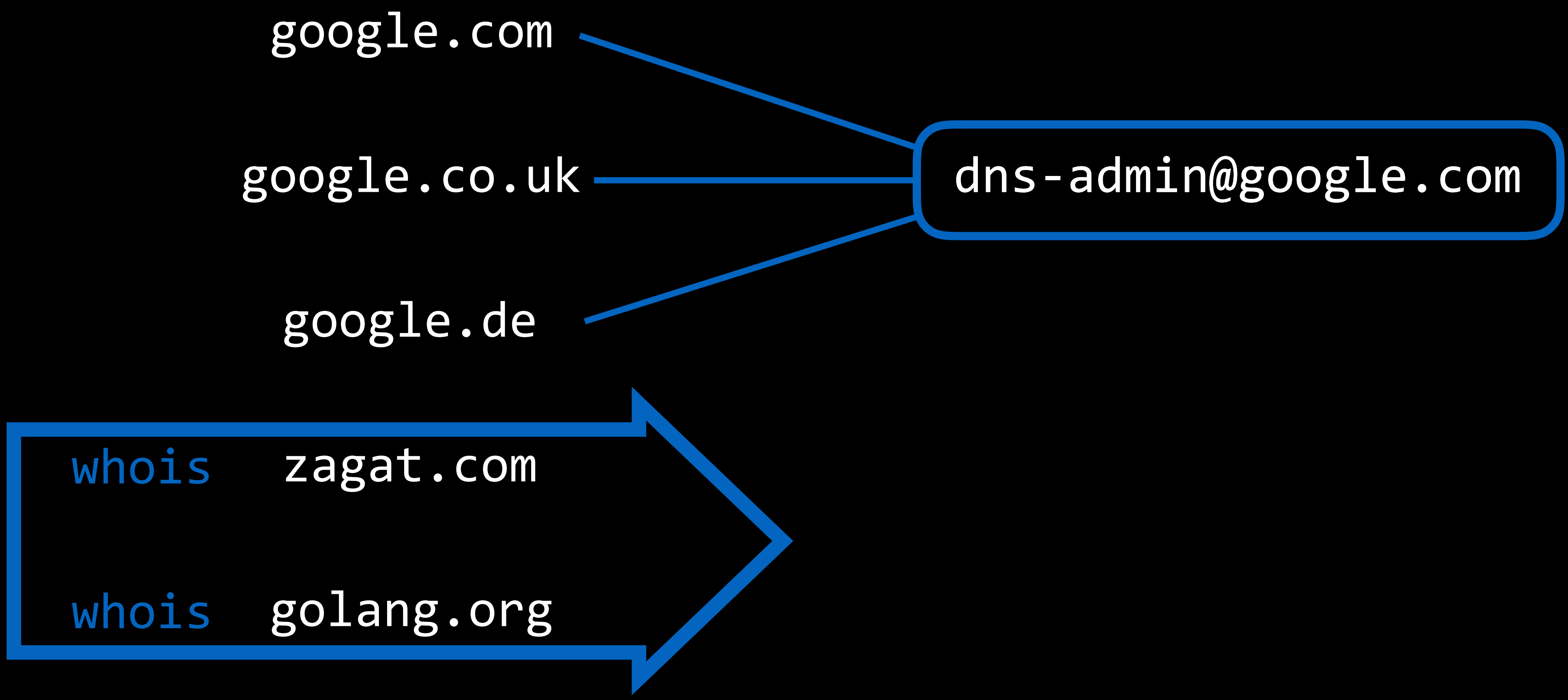
zagat.com

golang.org

dns-admin@google.com

Emails in whois records reflect administrative domain

Incorporating whois



google.com

google.co.uk

google.de

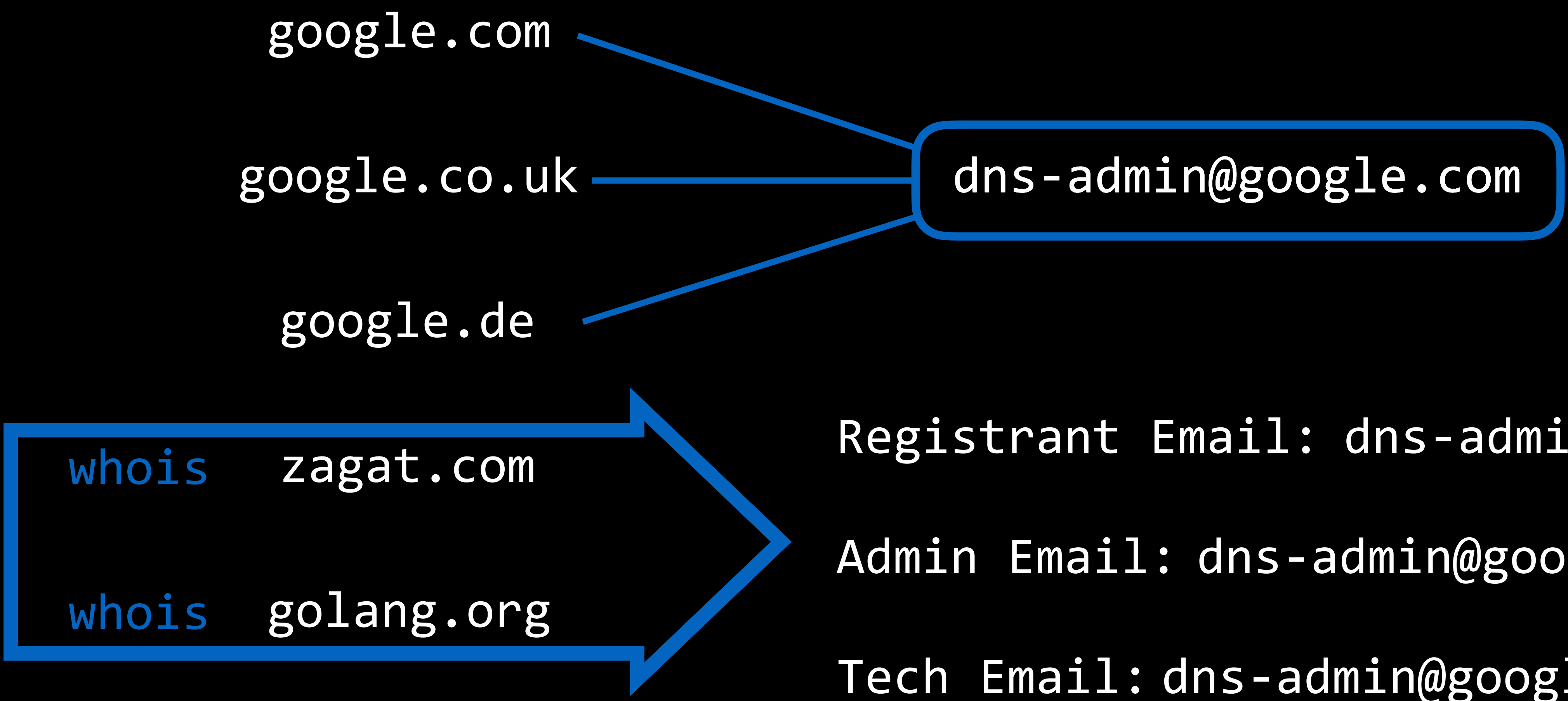
dns-admin@google.com

```
whois   zagat.com

whois  golang.org
```

Emails in whois records reflect administrative domain

Incorporating whois



google.com

google.co.uk

google.de

dns-admin@google.com

```
whois   zagat.com

whois  golang.org
```

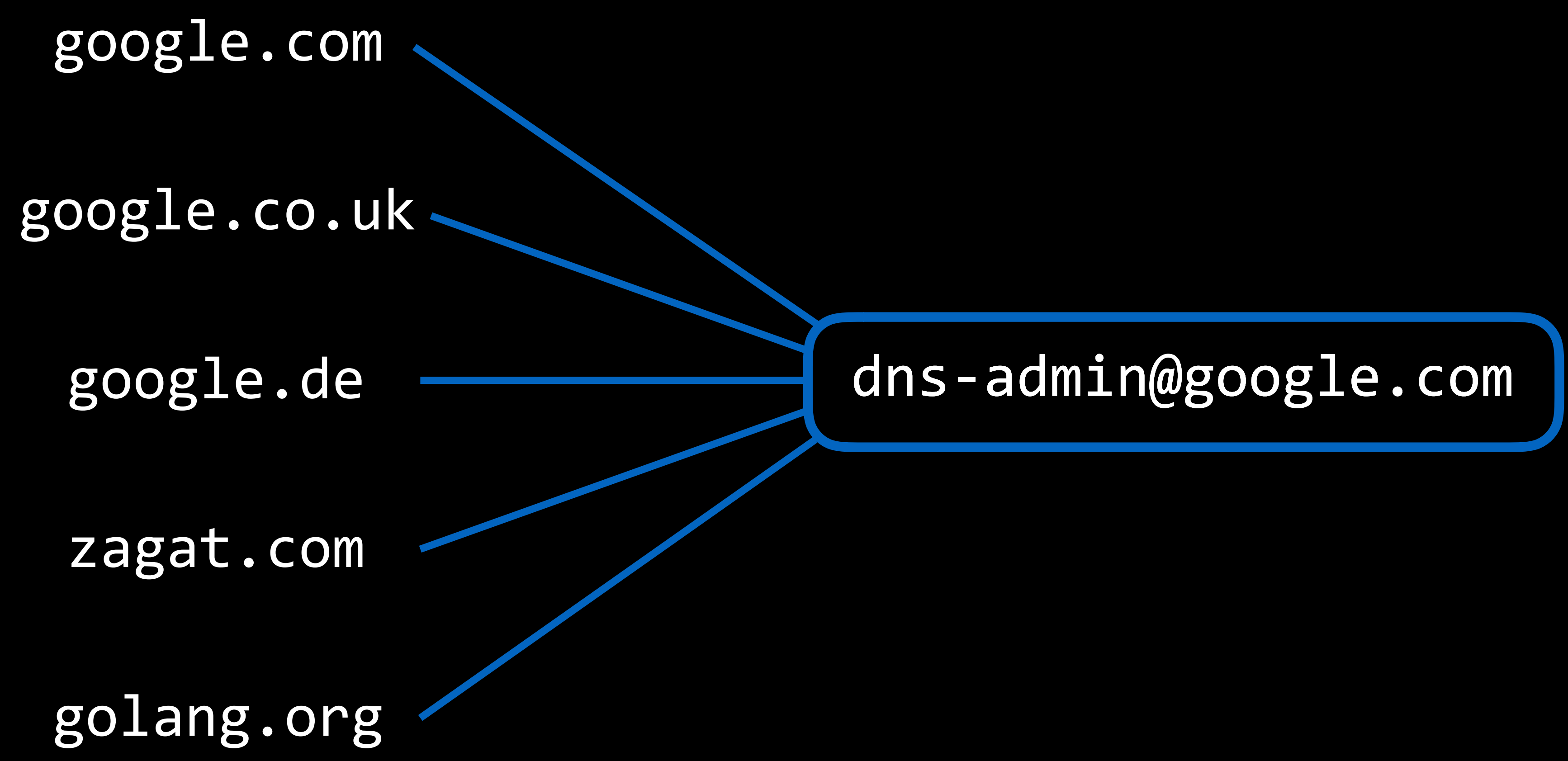
```
Registrant Email: dns-admin@google.com

Admin Email: dns-admin@google.com

Tech Email: dns-admin@google.com
```

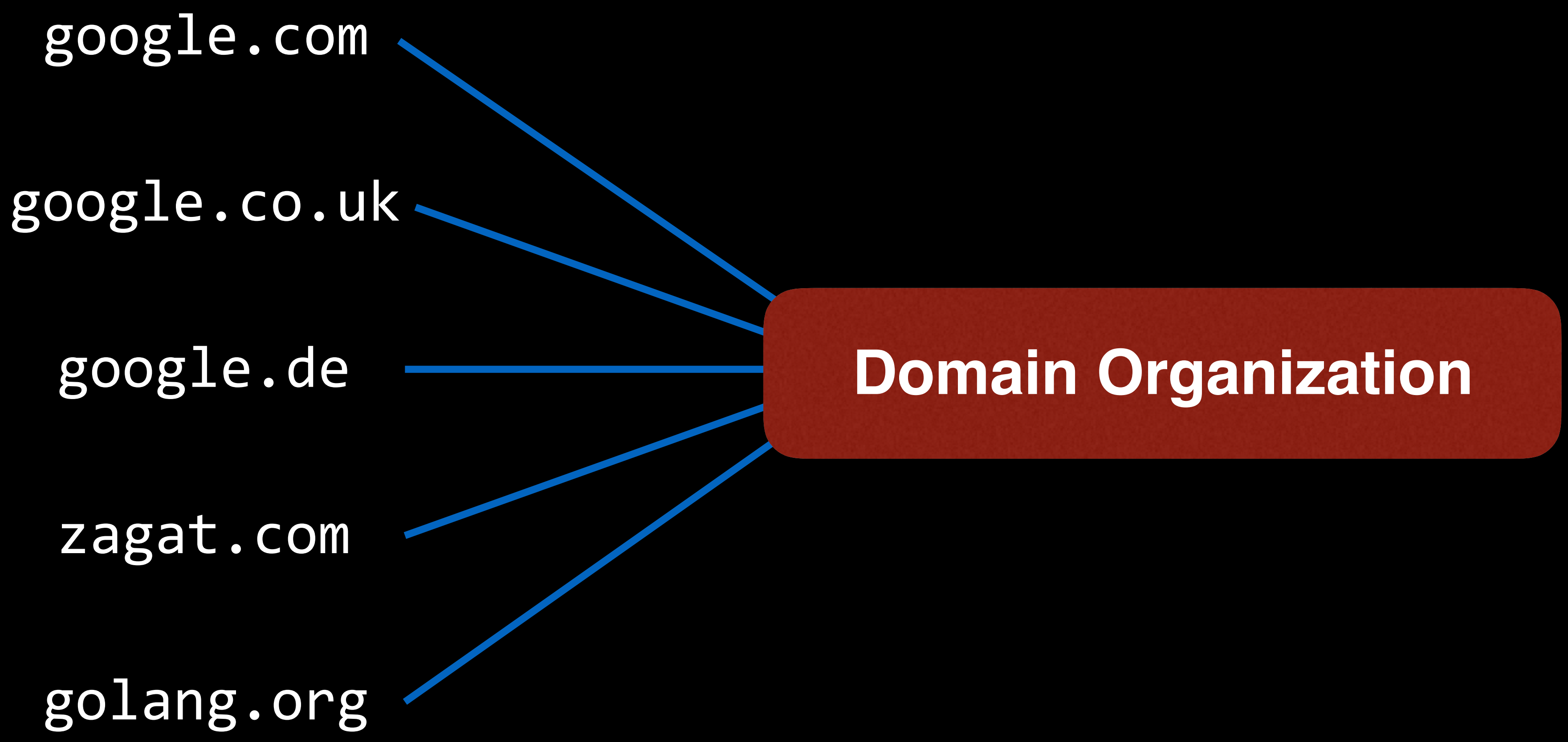
Emails in whois records reflect administrative domain

Incorporating whois



Emails in whois records reflect administrative domain

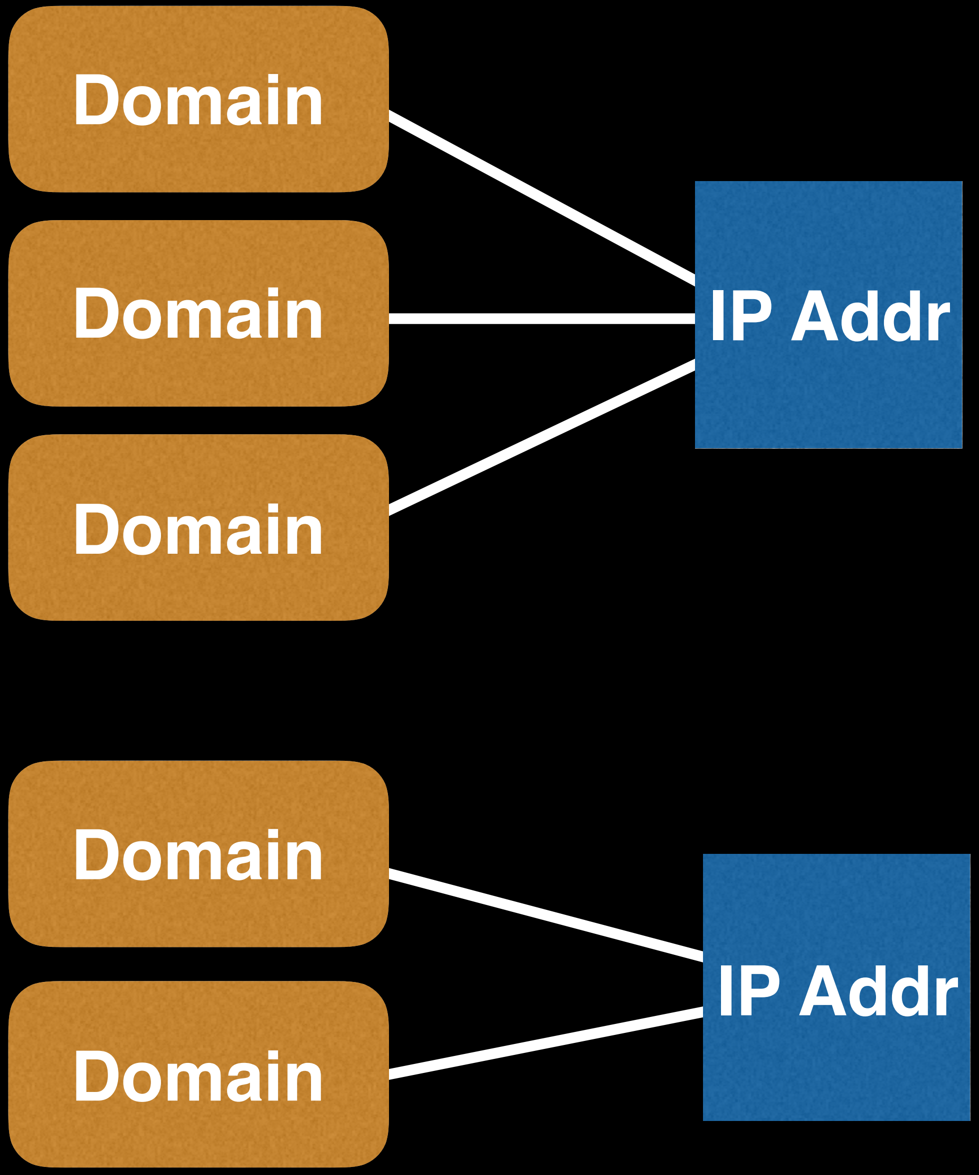
Incorporating whois



Emails in whois records reflect administrative domain

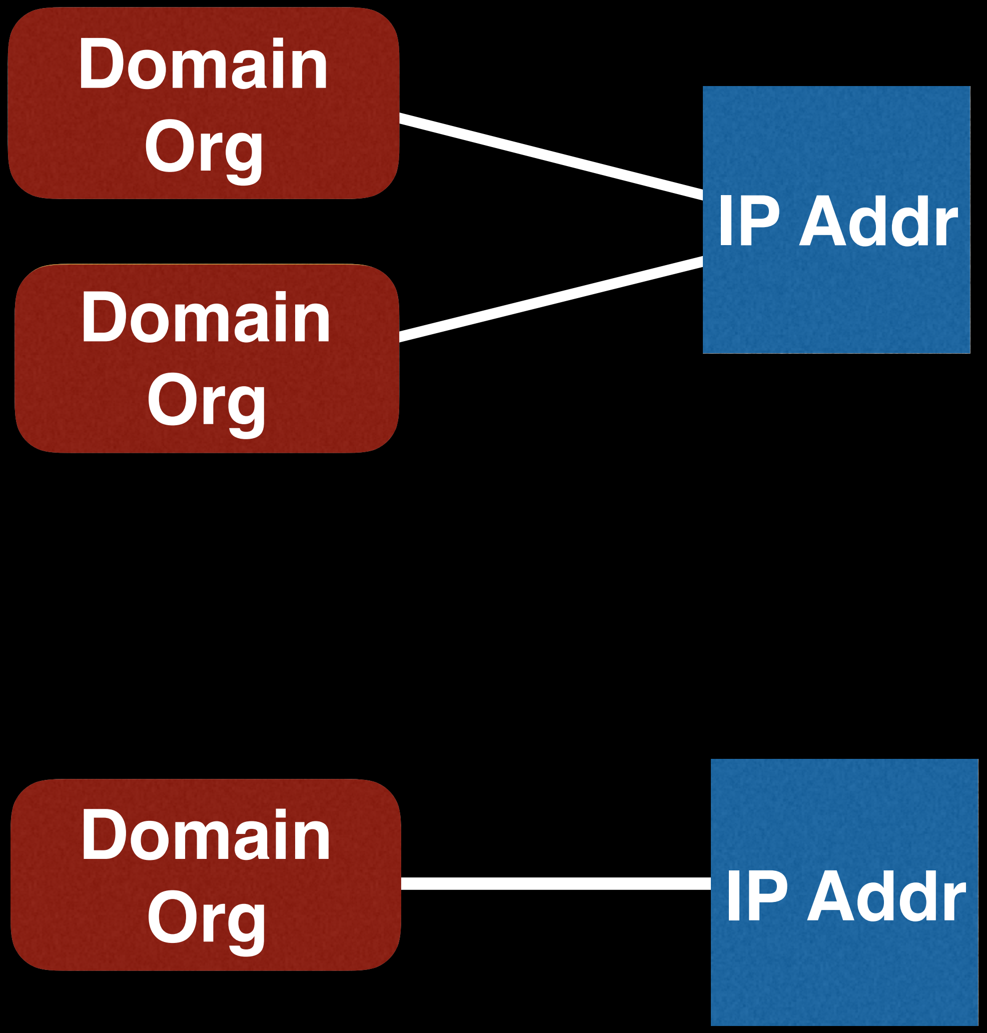
How do we detect sharing *at scale*?

Does the same entity that owns the domain own and operate the server at that IP address?



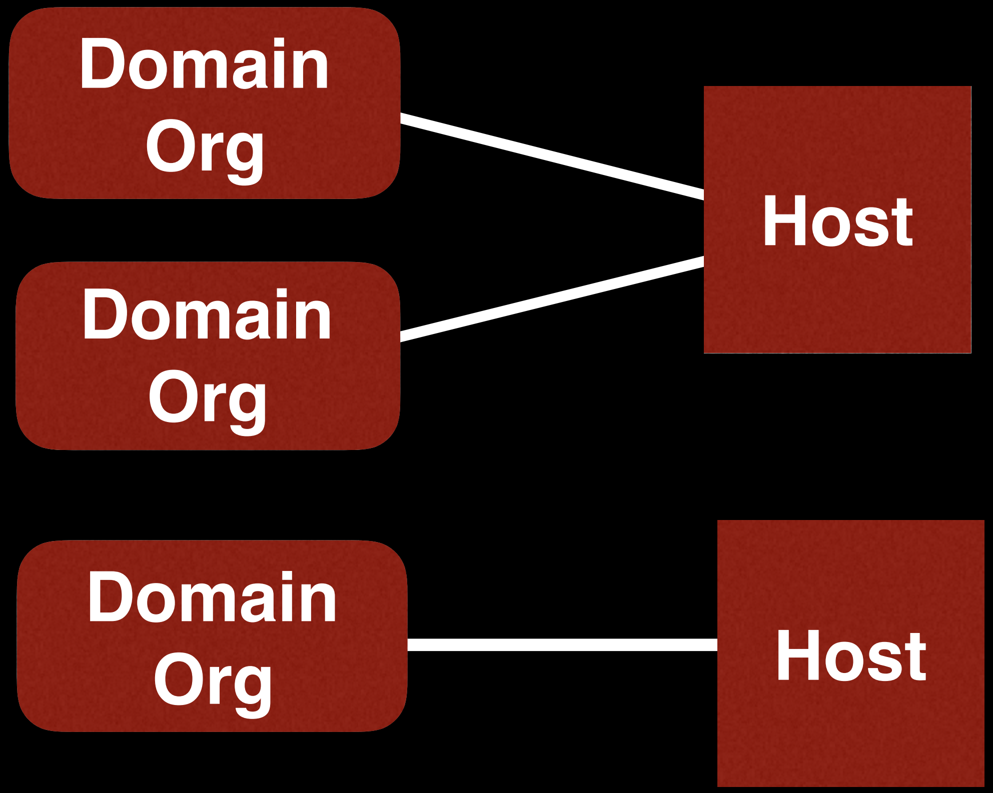
How do we detect sharing *at scale*?

Does the same entity that owns the domain own and operate the server at that IP address?



How do we detect sharing *at scale*?

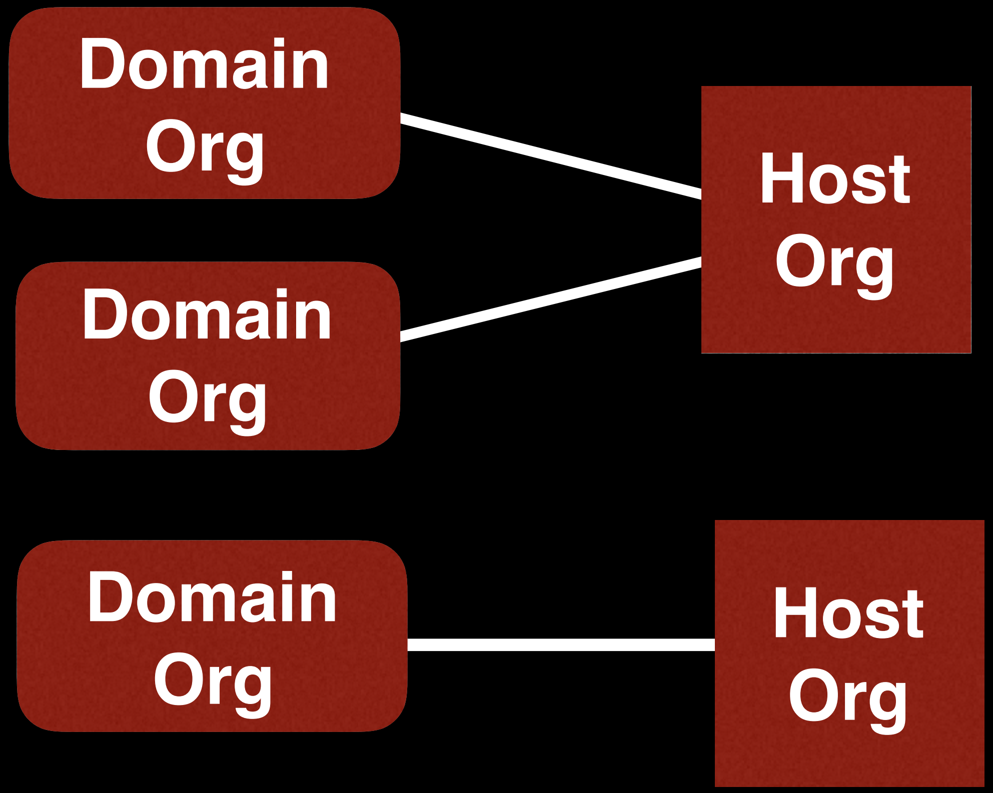
Does the same entity that owns the domain own and operate the server at that IP address?



Key sharing: domain org ≠ host org

How do we detect sharing *at scale*?

Does the same entity that owns the domain own and operate the server at that IP address?



Key sharing: domain org ≠ host org

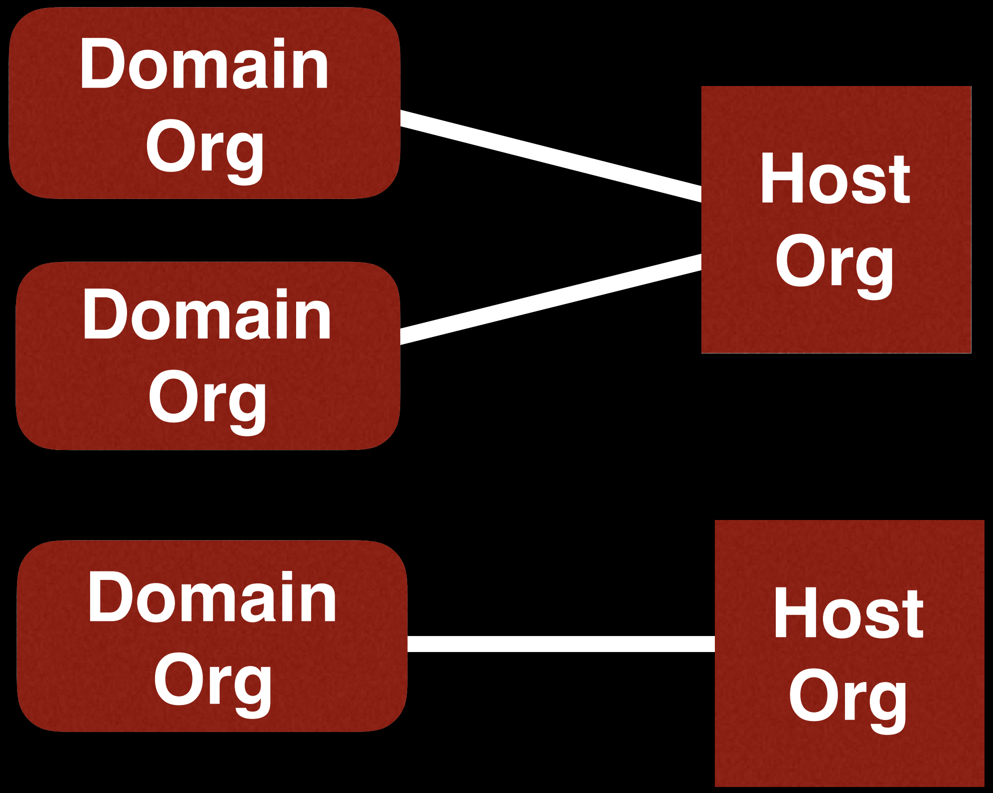
Outline

How prevalent is key sharing?

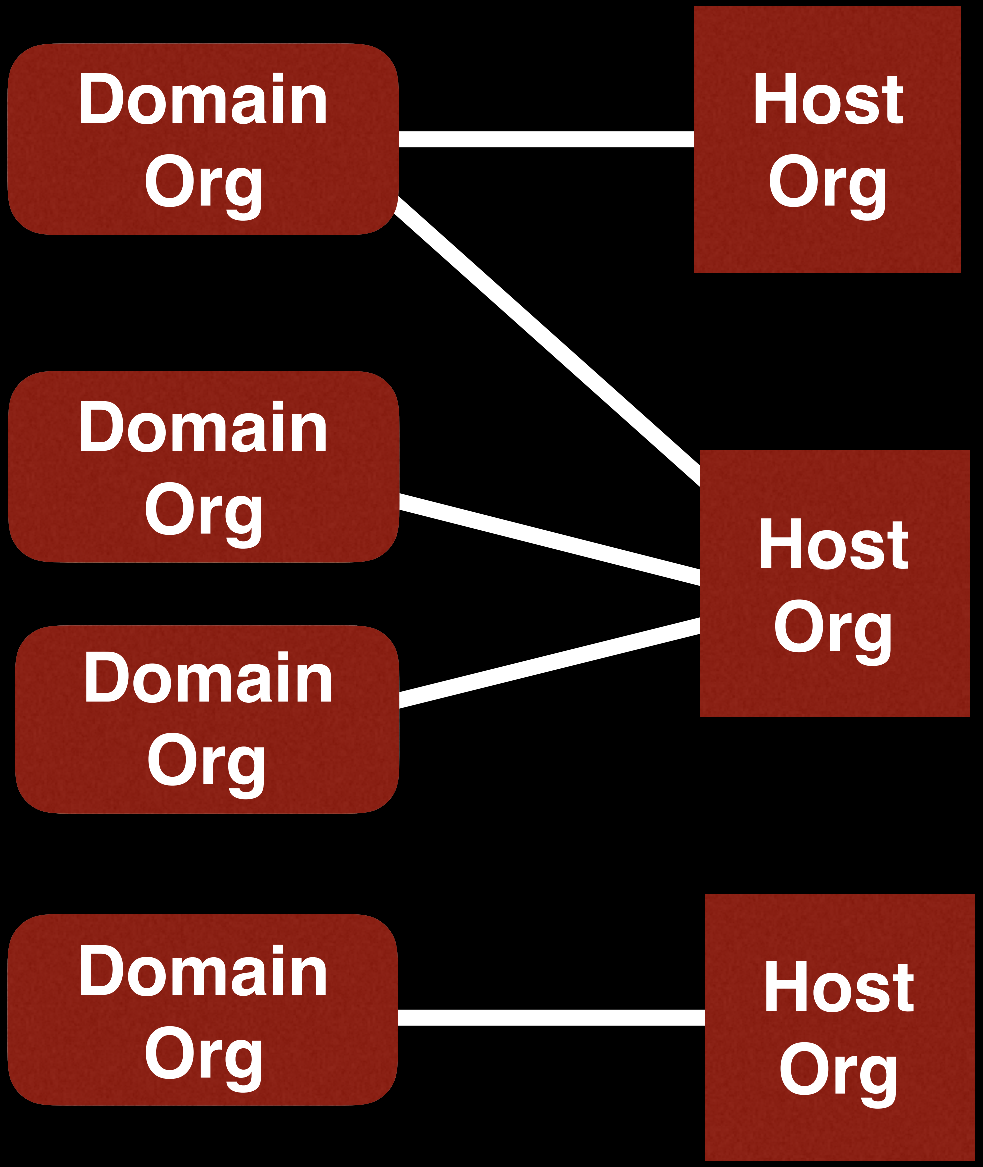
How many keys have providers aggregated?

How does sharing impact key management?

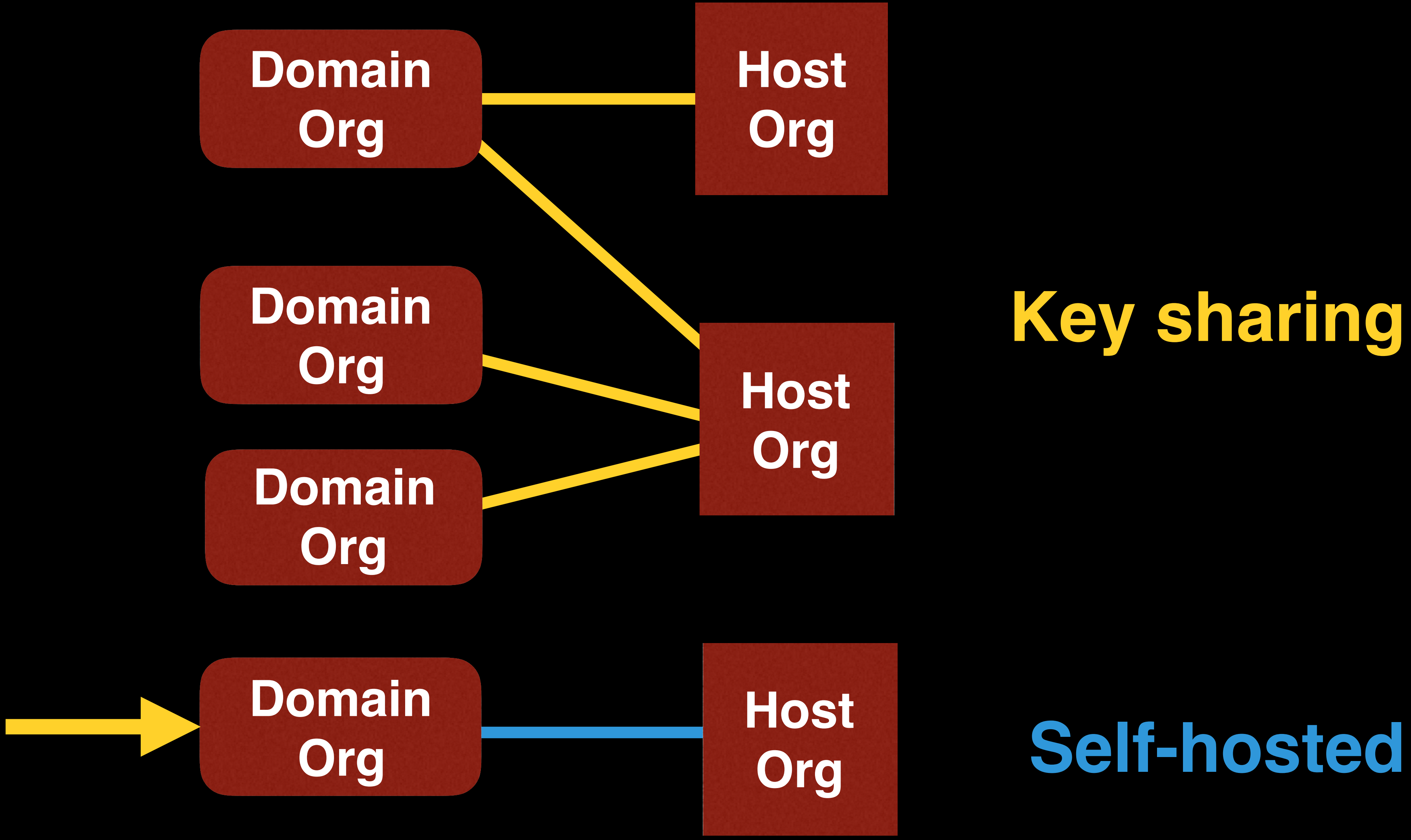
How prevalent is key sharing?



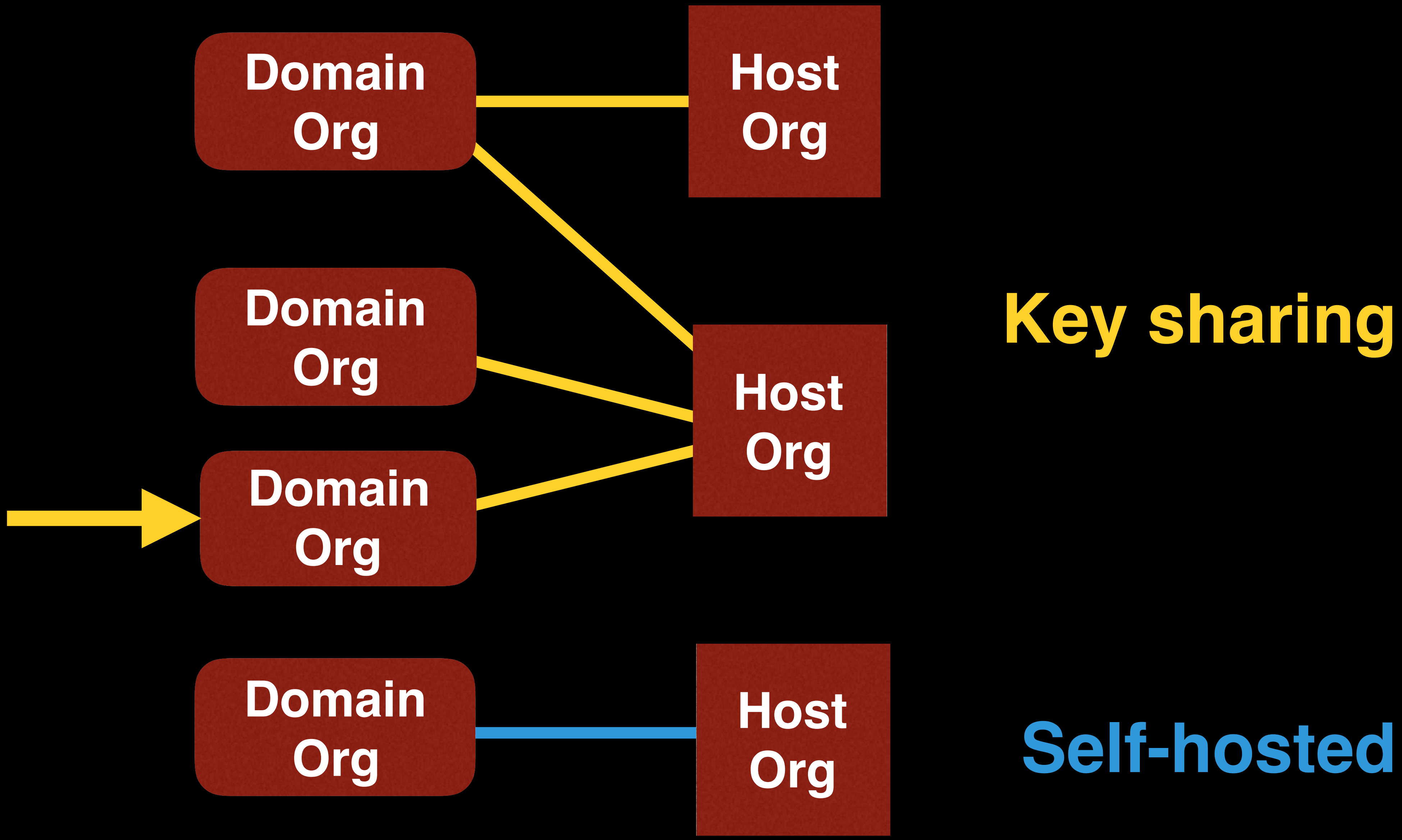
How prevalent is key sharing?



How prevalent is key sharing?



How prevalent is key sharing?



Domain Org

Host Org

Domain Org

Domain Org

Host Org

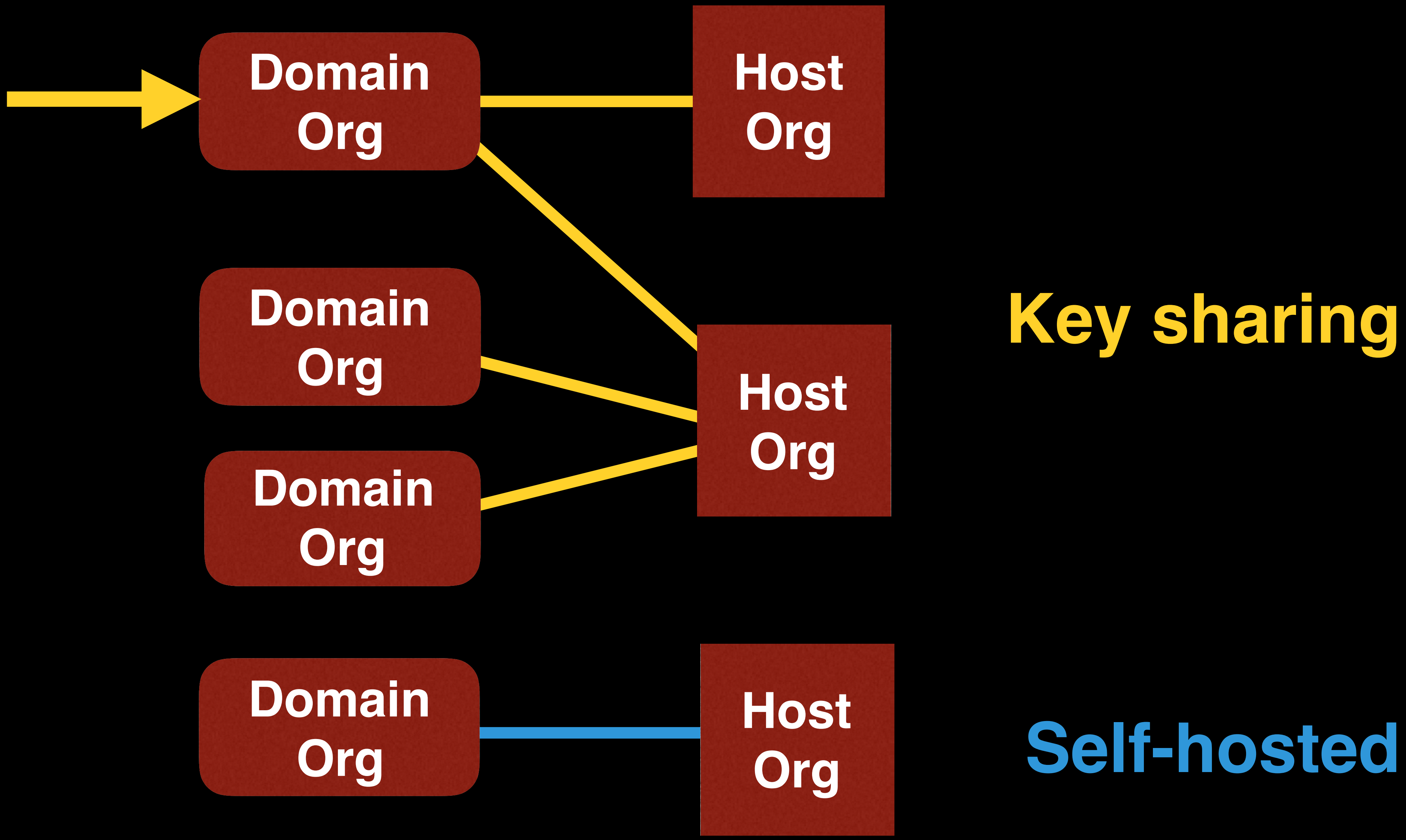
Key sharing

Domain Org

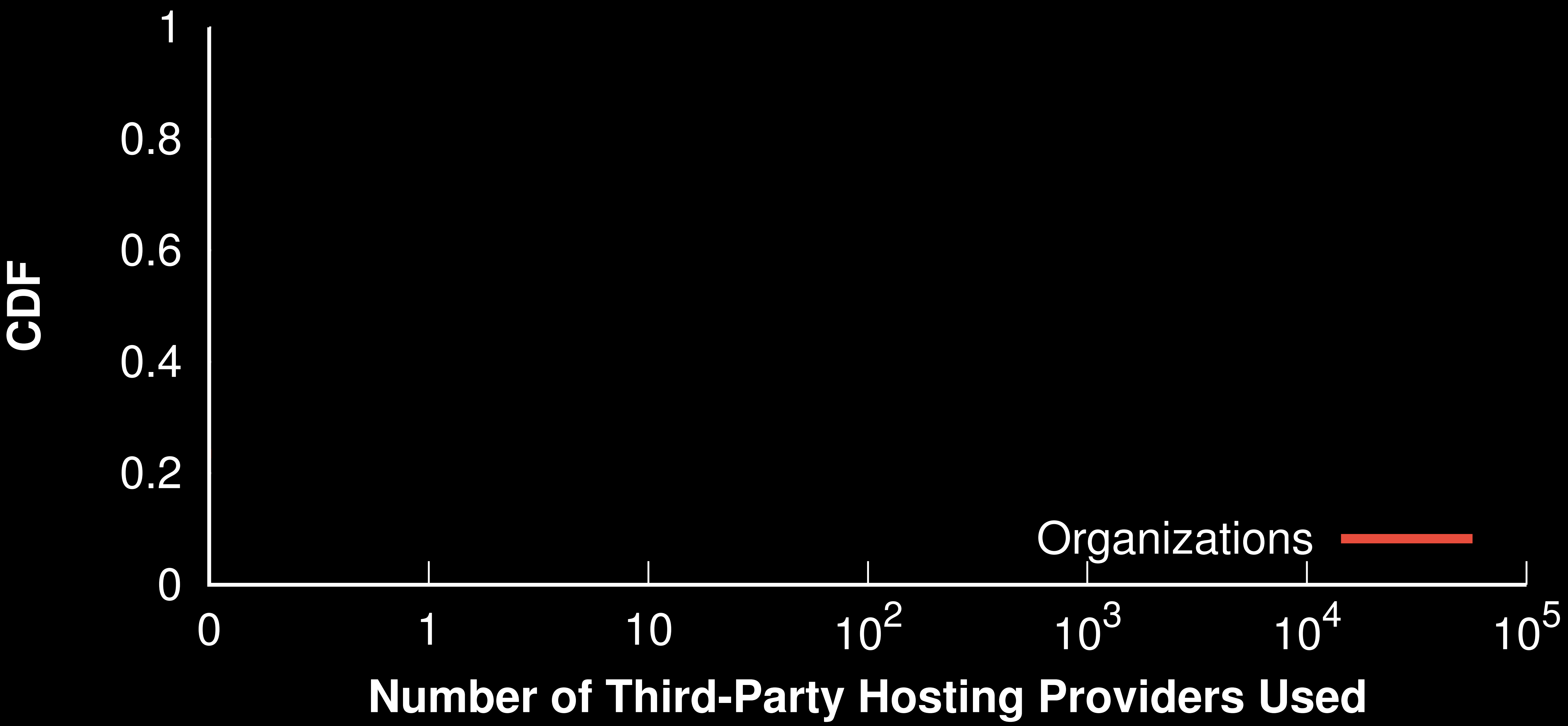
Host Org

Self-hosted

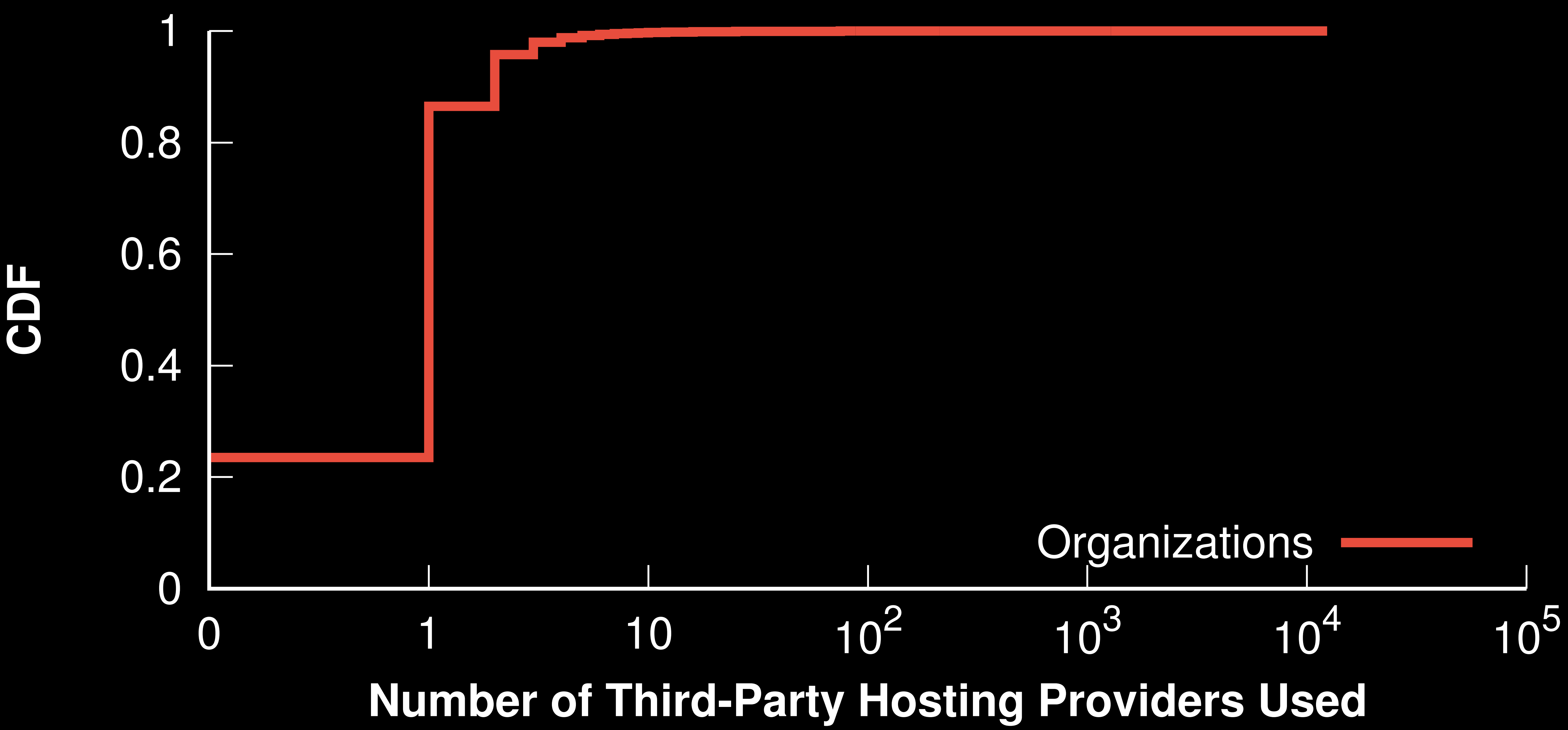
How prevalent is key sharing?



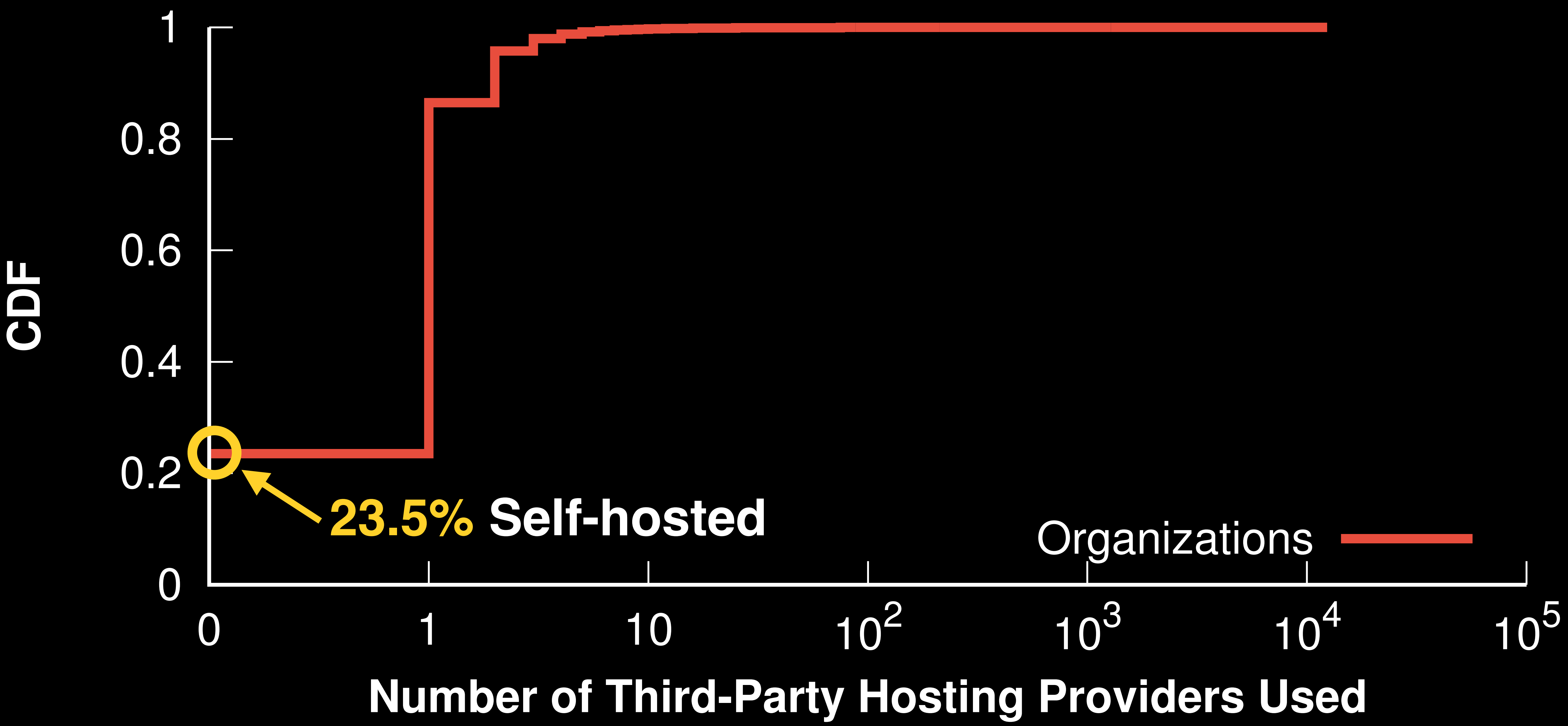
How prevalent is key sharing?



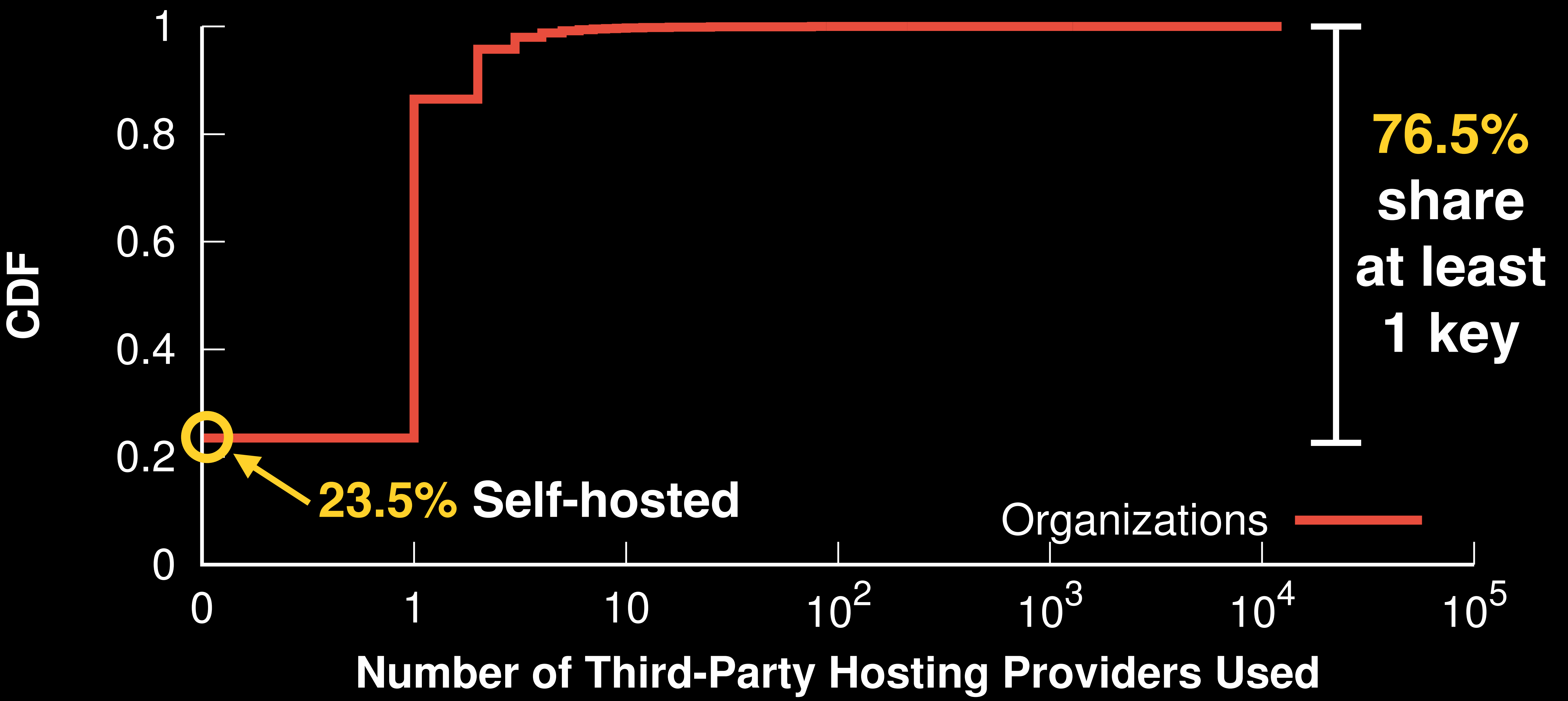
How prevalent is key sharing?



How prevalent is key sharing?



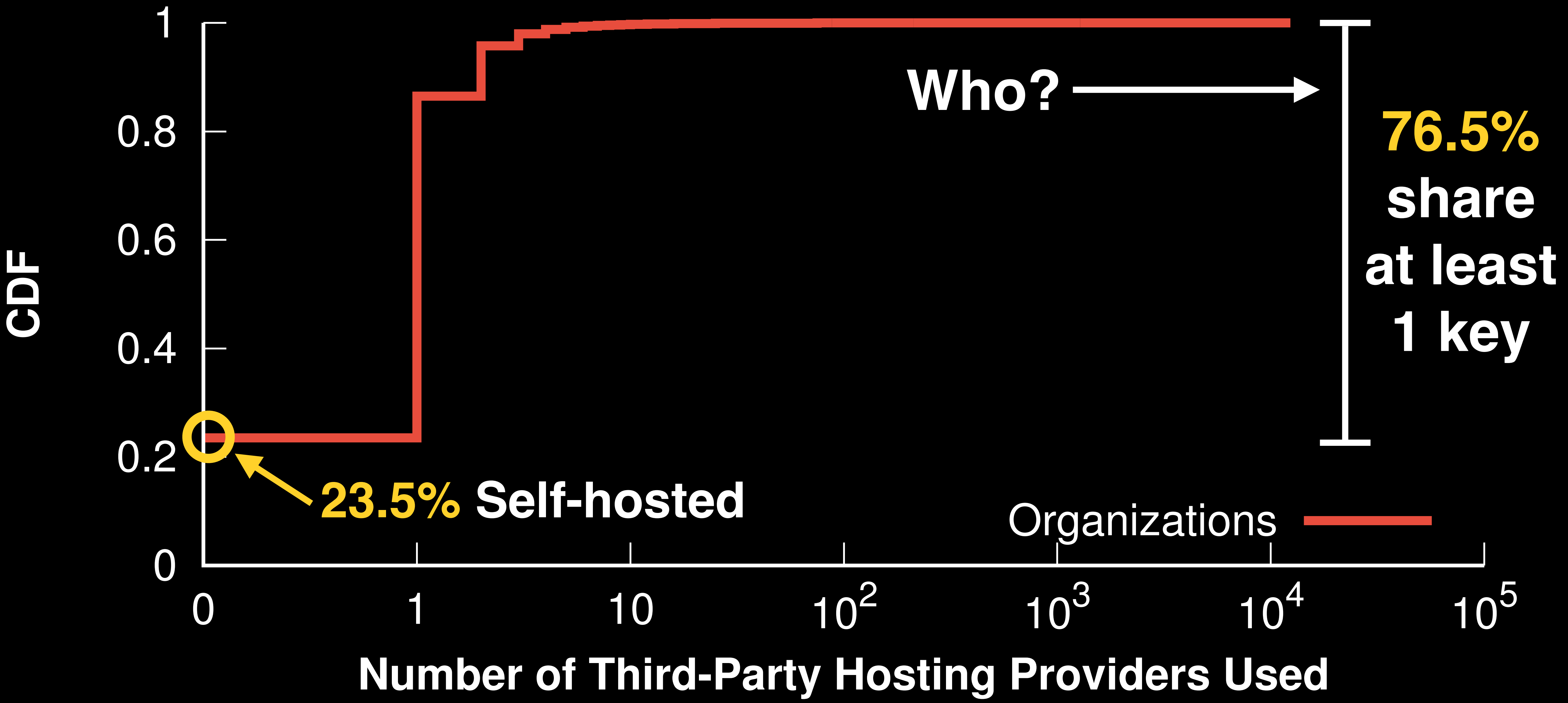
How prevalent is key sharing?



23.5% Self-hosted

76.5% share at least 1 key

How prevalent is key sharing?

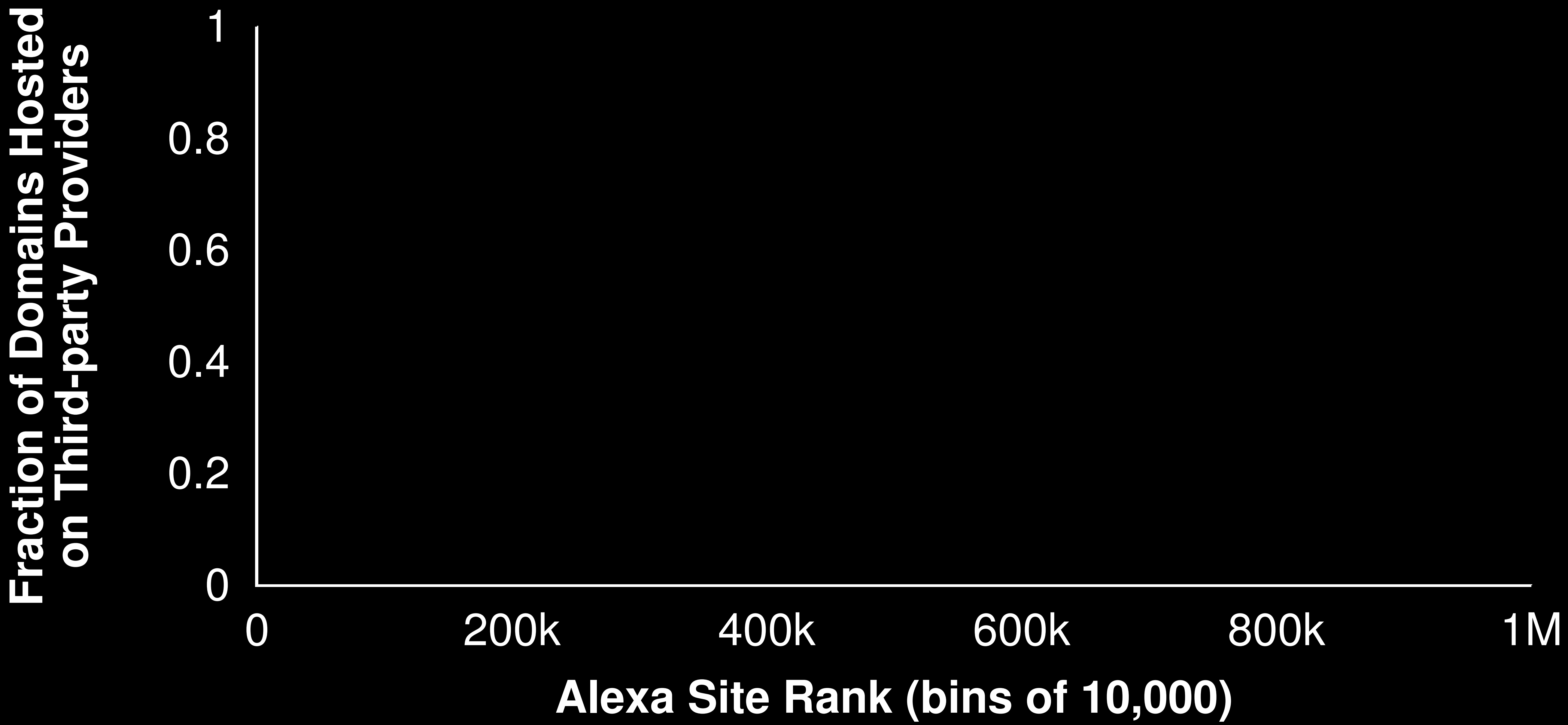


Who?

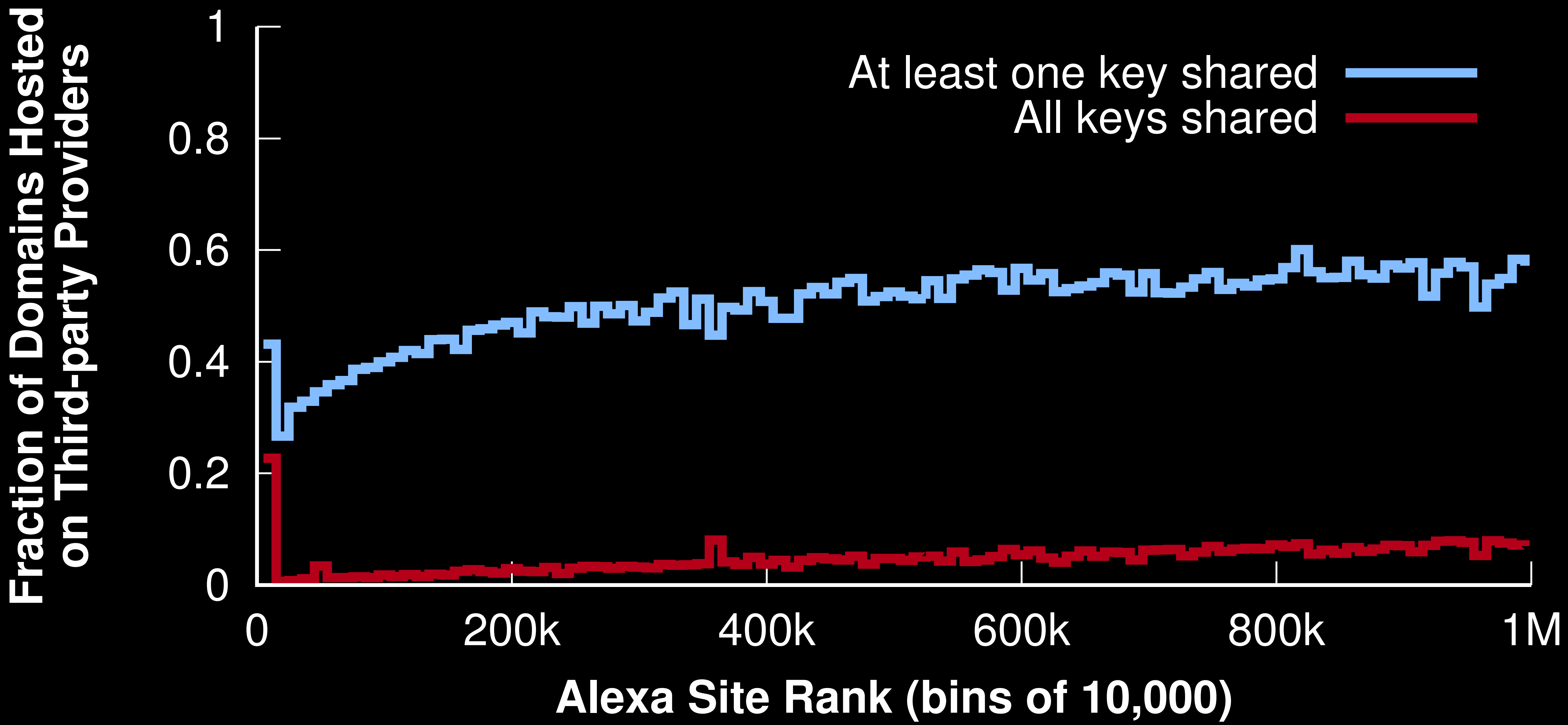
76.5% share at least 1 key

23.5% Self-hosted

Who shares?

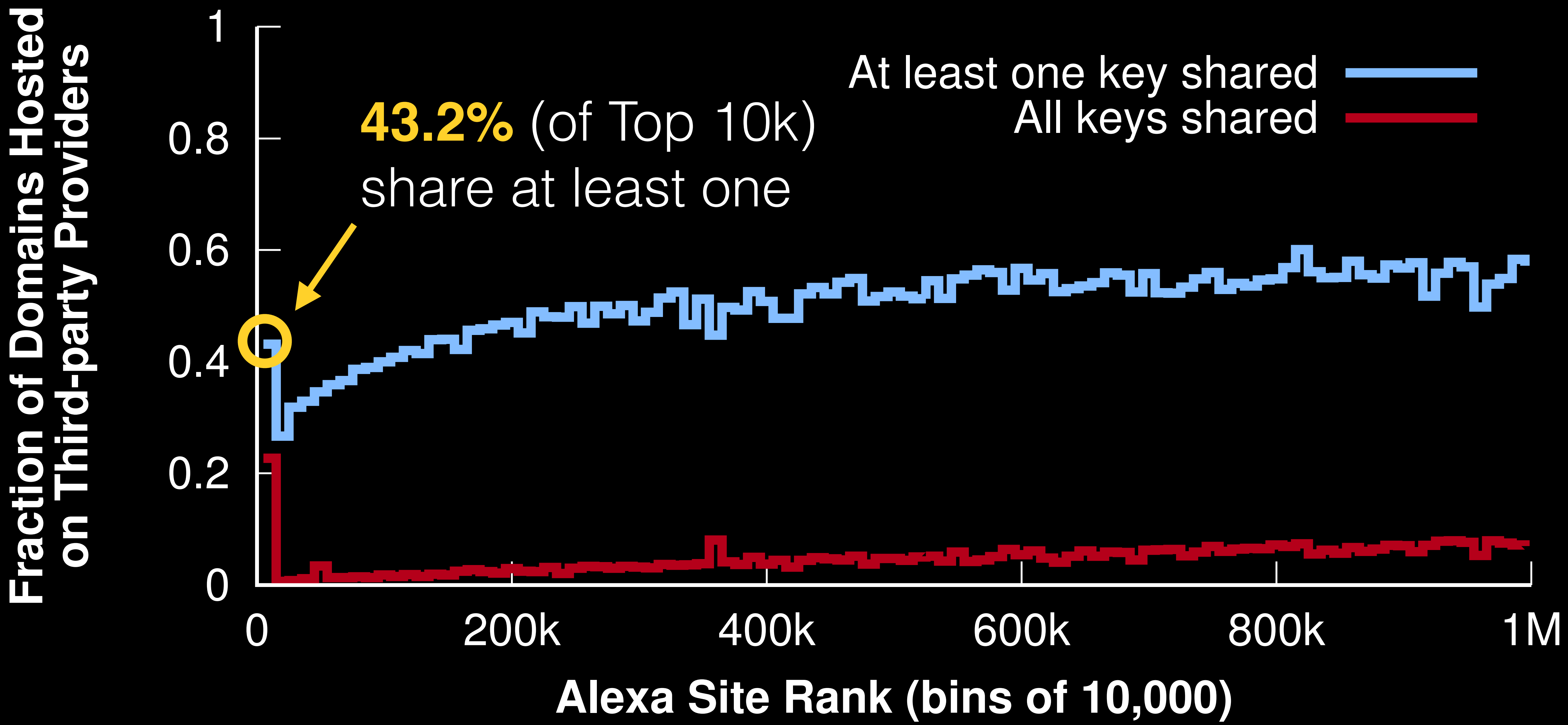

Fraction of Domains Hosted on Third-party Providers
1
0.8
0.6
0.4
0.2
0
0
200k
400k
600k
800k
1M
Alexa Site Rank (bins of 10,000)

Who shares?



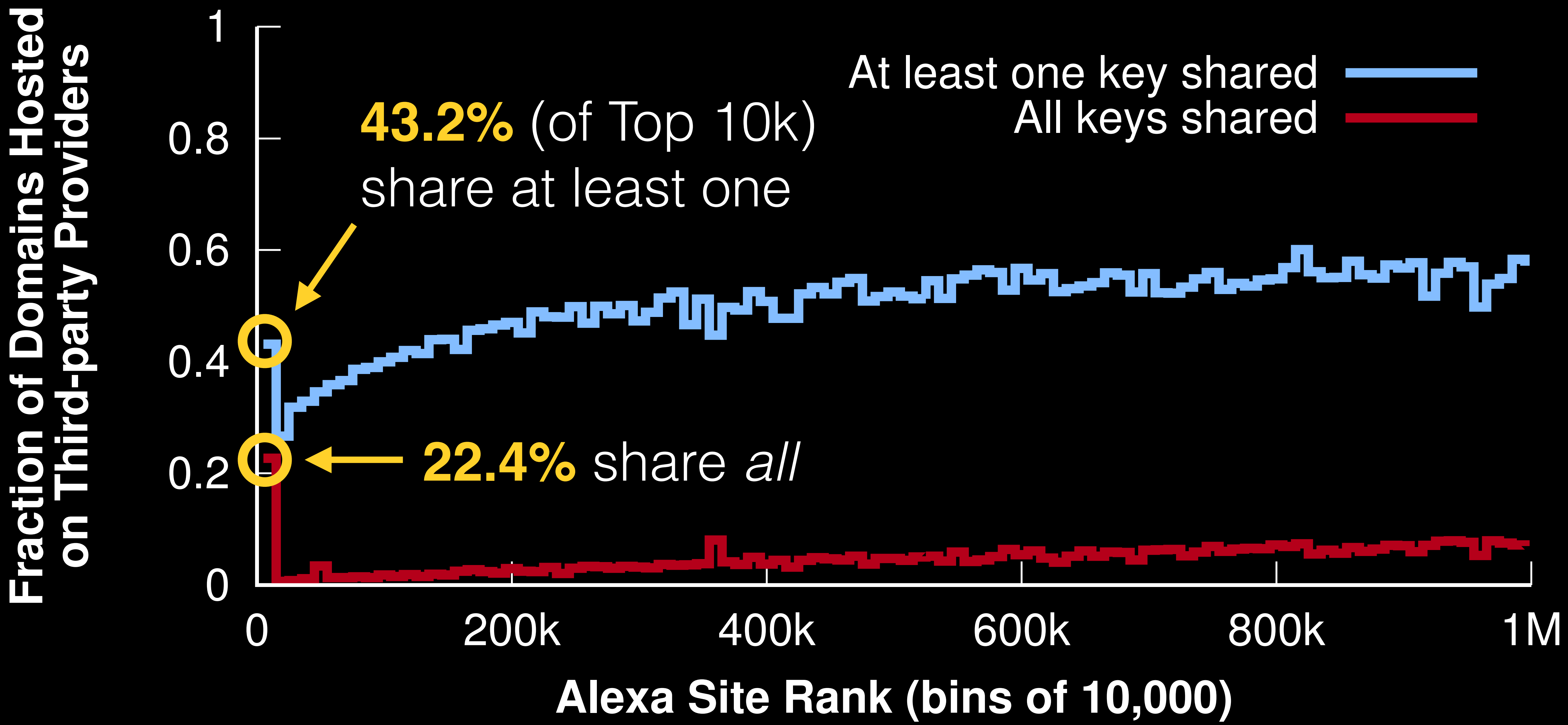
Fraction of Domains Hosted on Third-party Providers
1
0.8
0.6
0.4
0.2
0
0
200k
400k
600k
800k
1M
Alexa Site Rank (bins of 10,000)
At least one key shared
All keys shared

Who shares?



43.2% (of Top 10k) share at least one

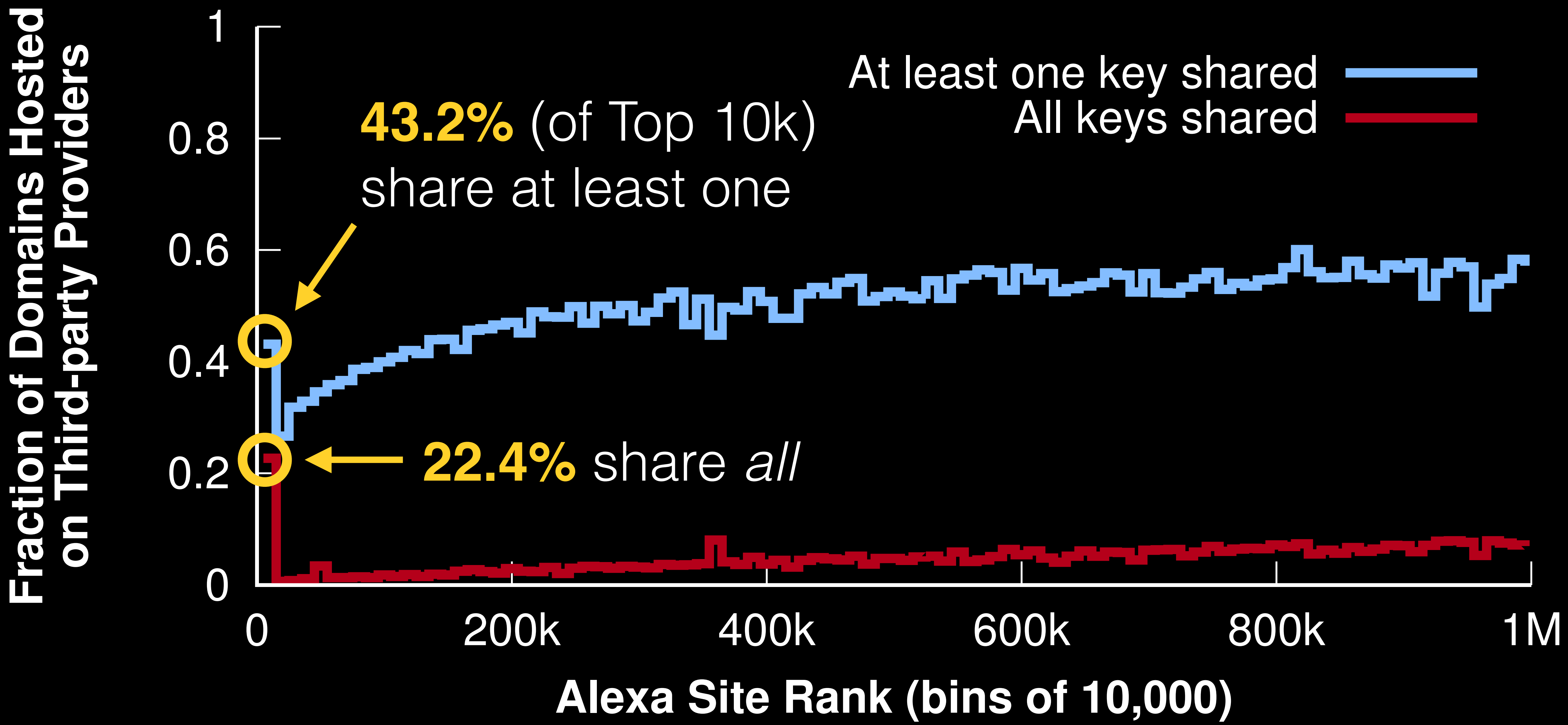
Who shares?



43.2% (of Top 10k) share at least one

22.4% share *all*

Who shares?



43.2% (of Top 10k) share at least one

22.4% share *all*

Key sharing is common across the Internet

Outline

How prevalent is key sharing?

How many keys have providers aggregated?

How does sharing impact key management?

Outline

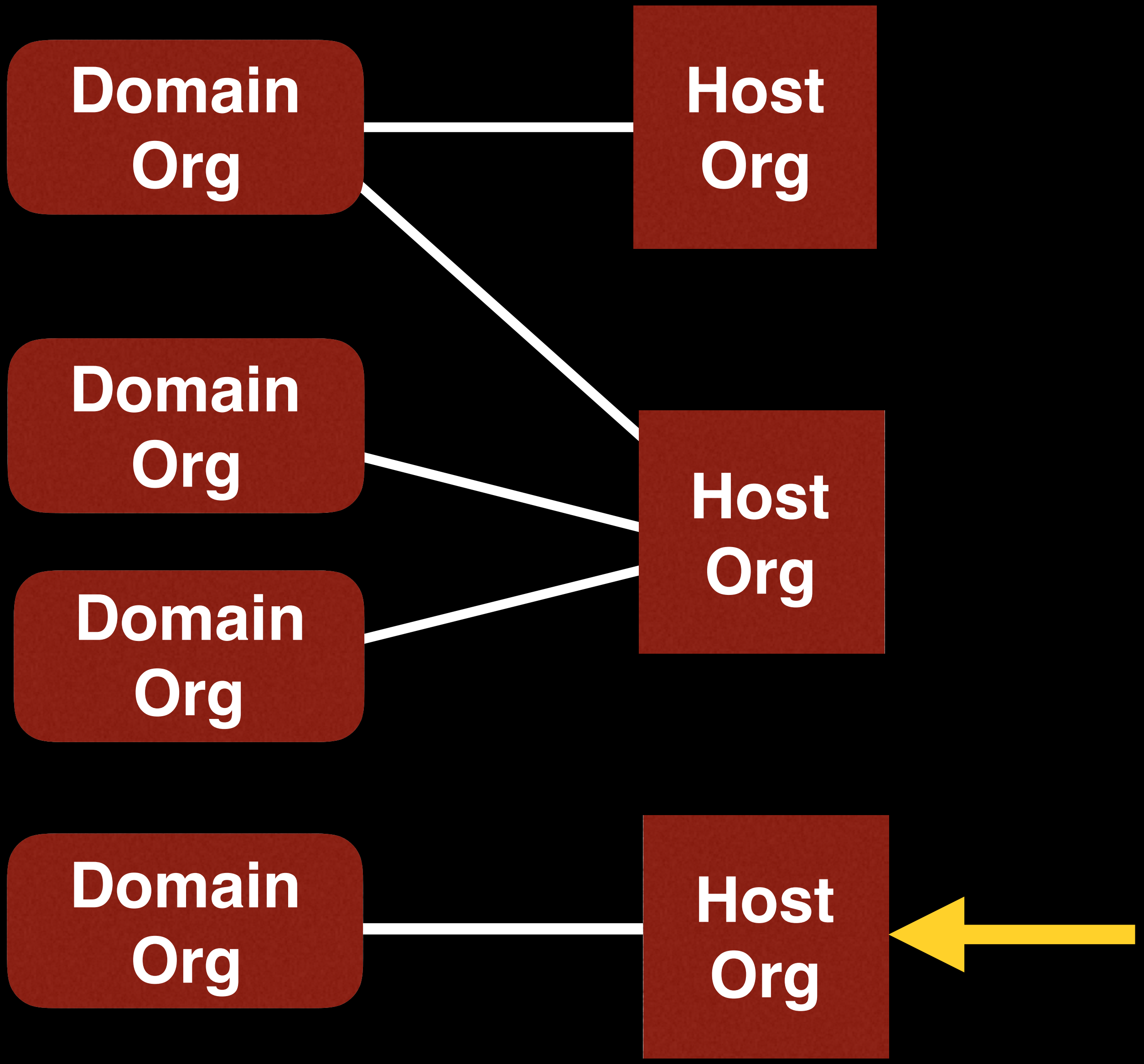
How prevalent is key sharing?

- 76.5% share with ≥ 1 provider
- Common even among most popular websites

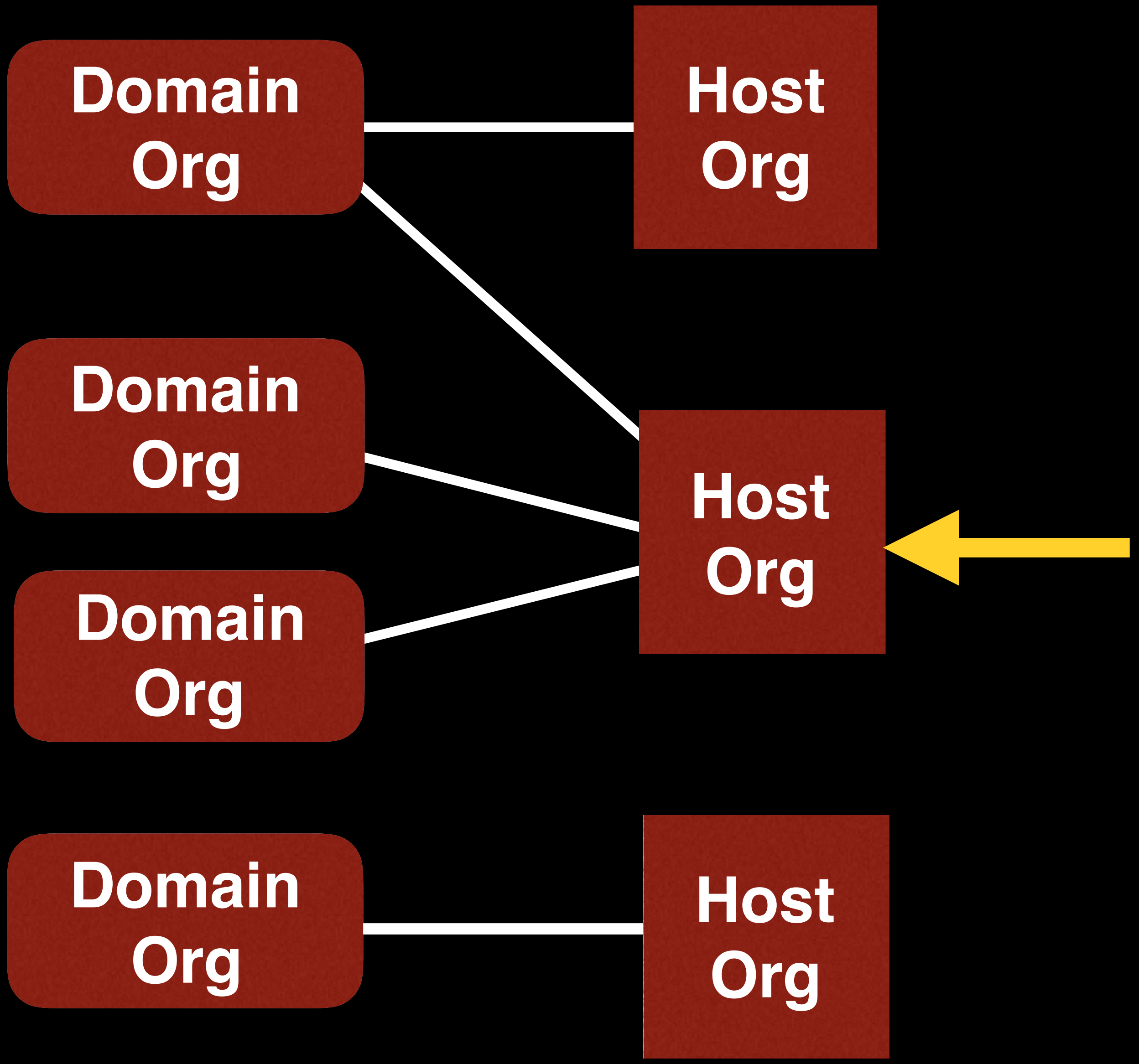
How many keys have providers aggregated?

How does sharing impact key management?

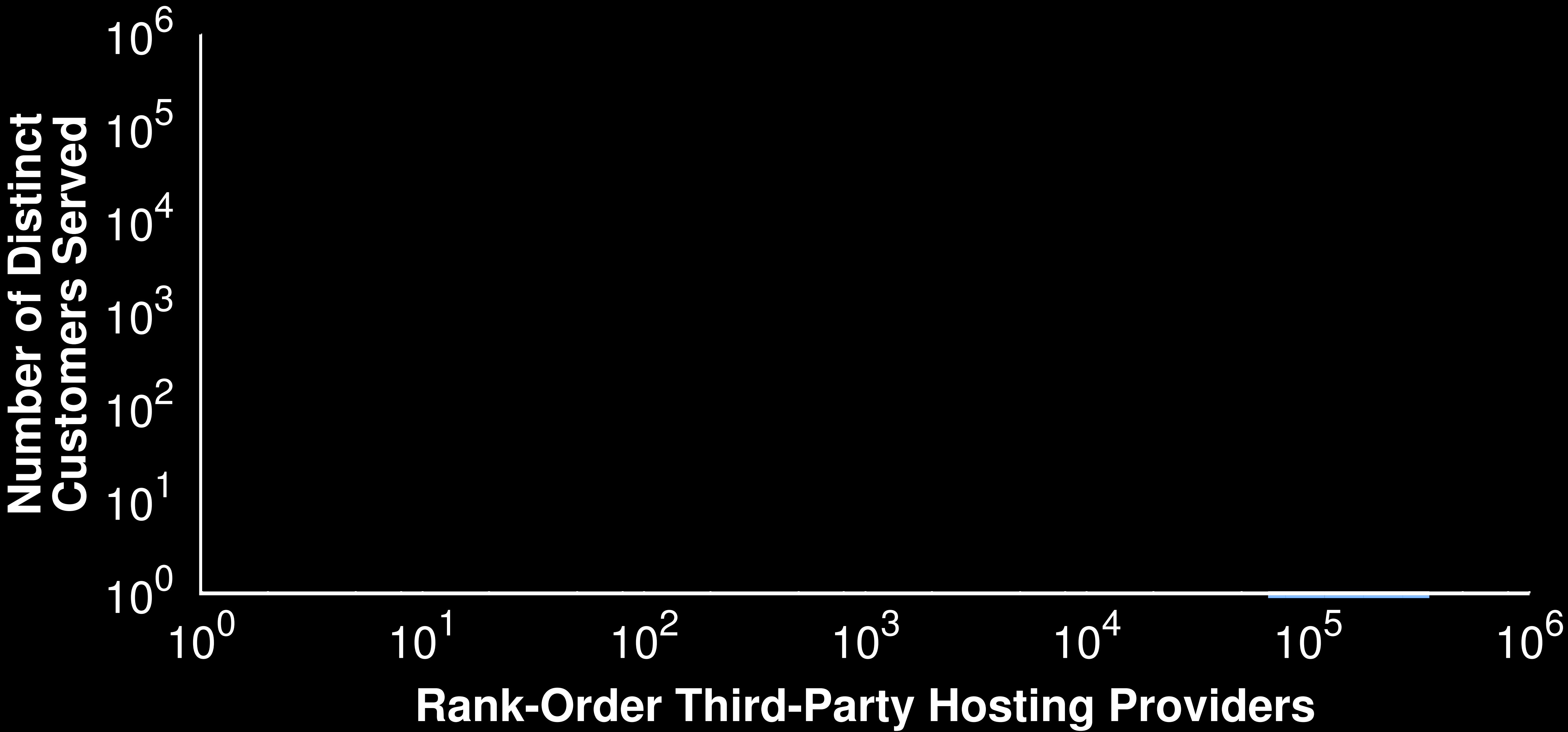
How many keys have providers aggregated?



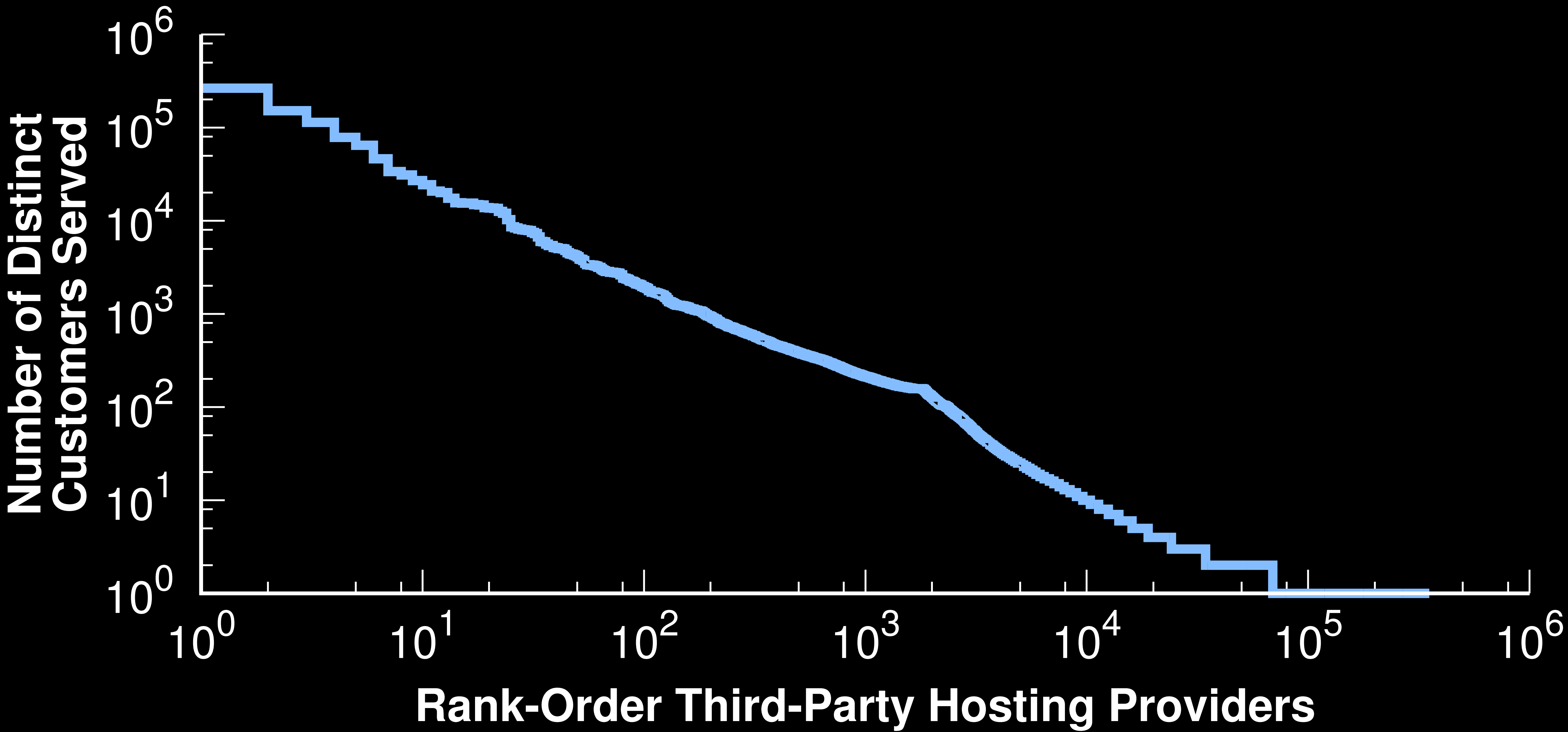
How many keys have providers aggregated?



How have keys been aggregated?

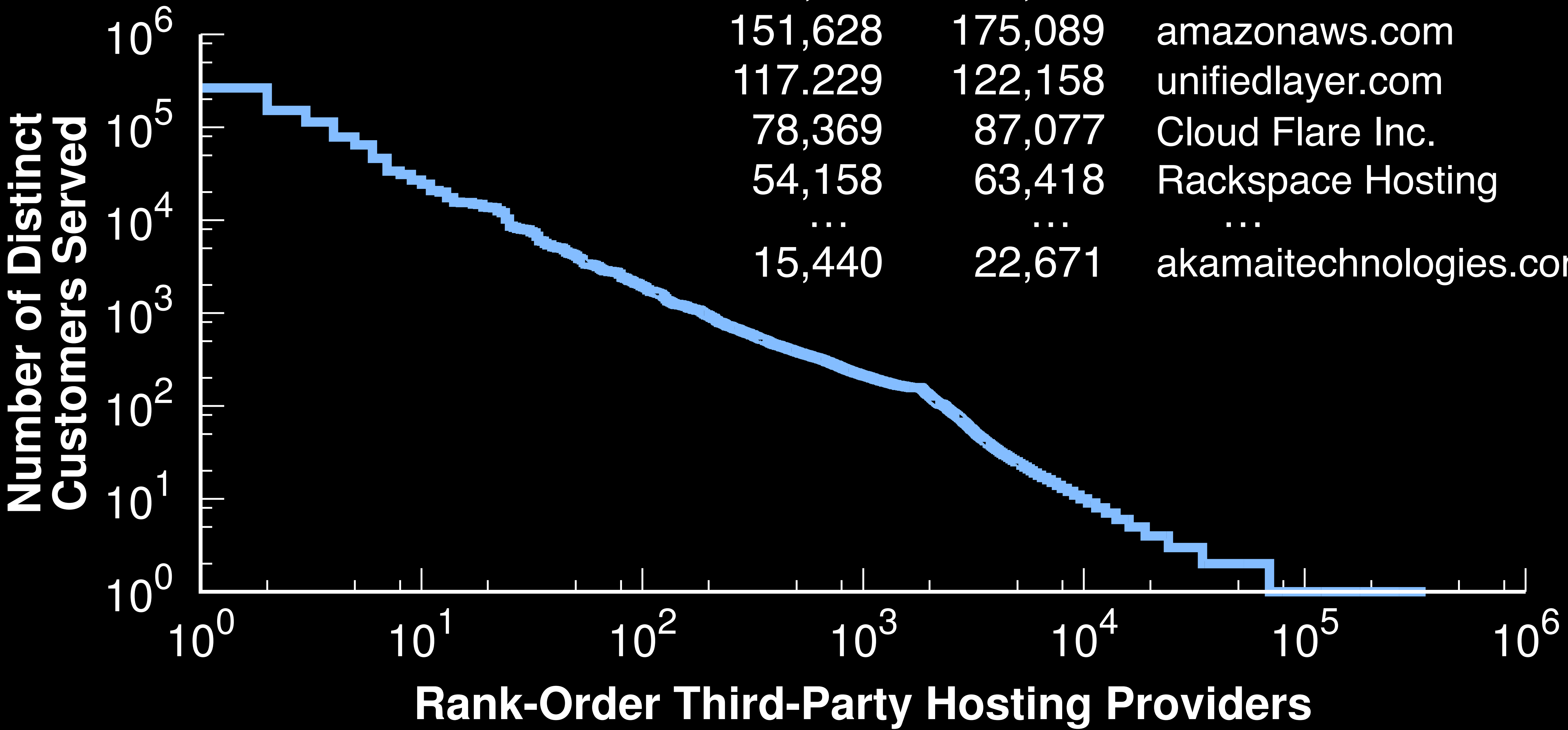


How have keys been aggregated?

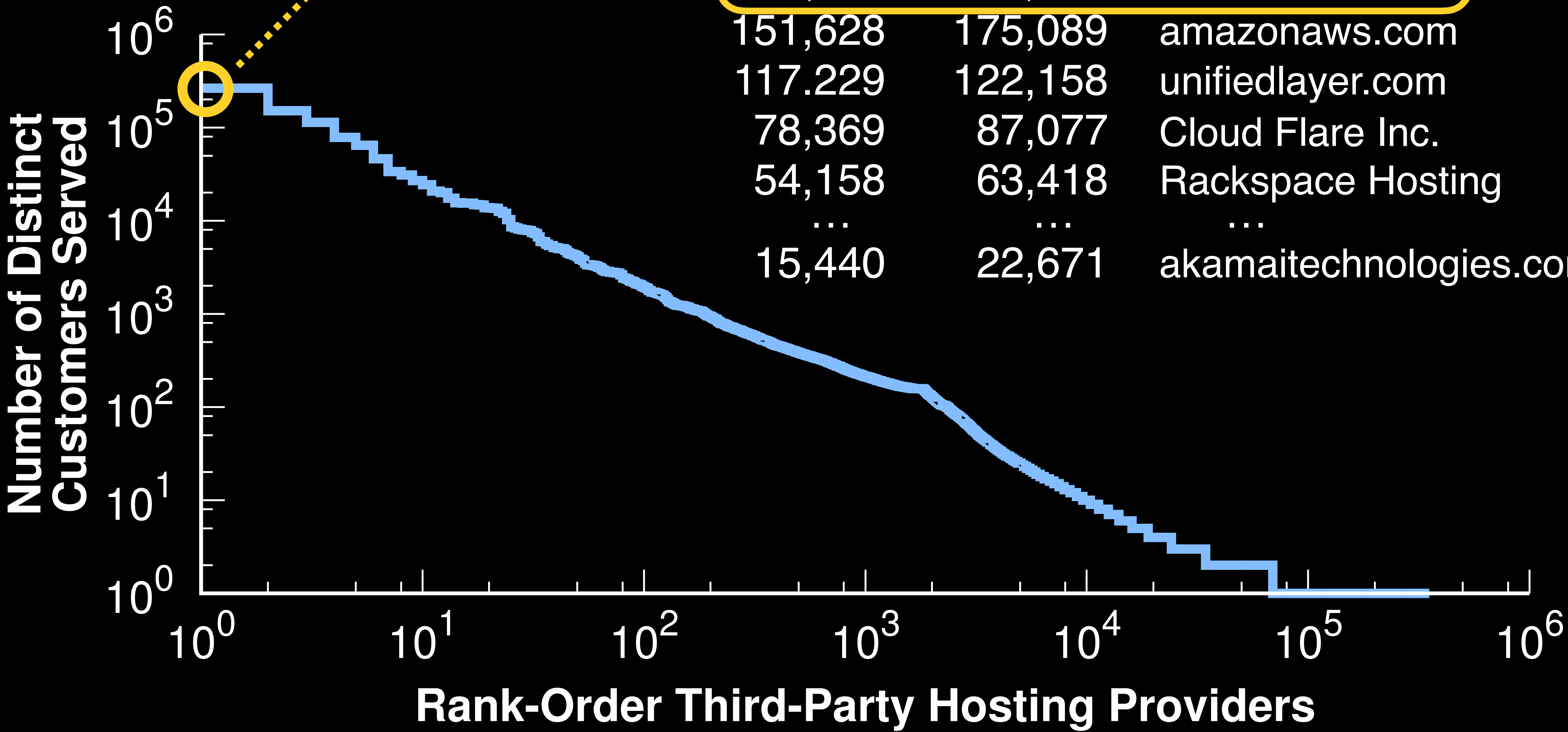


How have keys been aggregated?

#Organizations	#Domains	Hosting provider
266,110	277,891	secureserver.net
151,628	175,089	amazonaws.com
117.229	122,158	unifiedlayer.com
78,369	87,077	Cloud Flare Inc.
54,158	63,418	Rackspace Hosting
…	…	…
15,440	22,671	akamaitechnologies.com

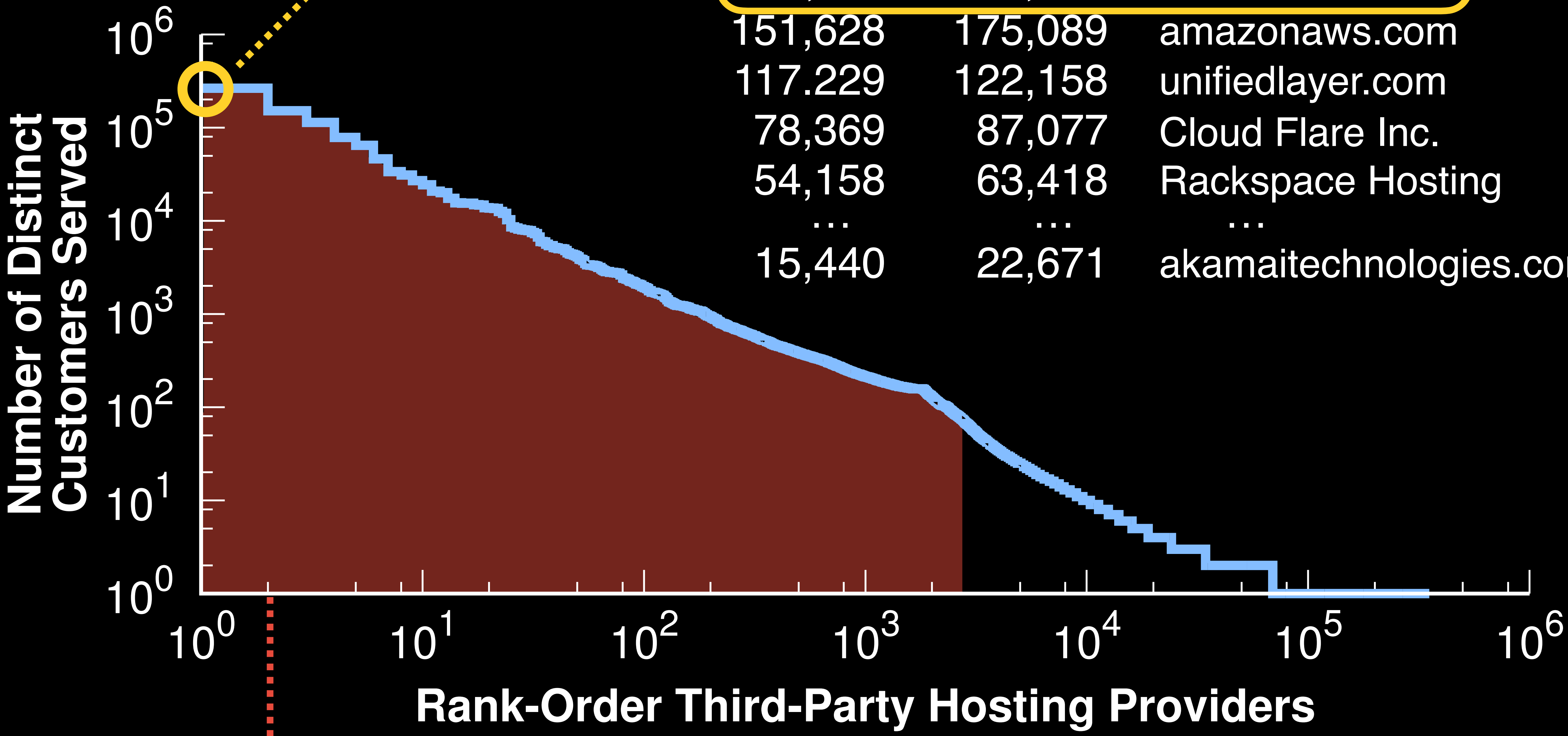


How have keys been aggregated?



#Organizations	#Domains	Hosting provider
266,110	277,891	secureserver.net
151,628	175,089	amazonaws.com
117.229	122,158	unifiedlayer.com
78,369	87,077	Cloud Flare Inc.
54,158	63,418	Rackspace Hosting
…	…	…
15,440	22,671	akamaitechnologies.com

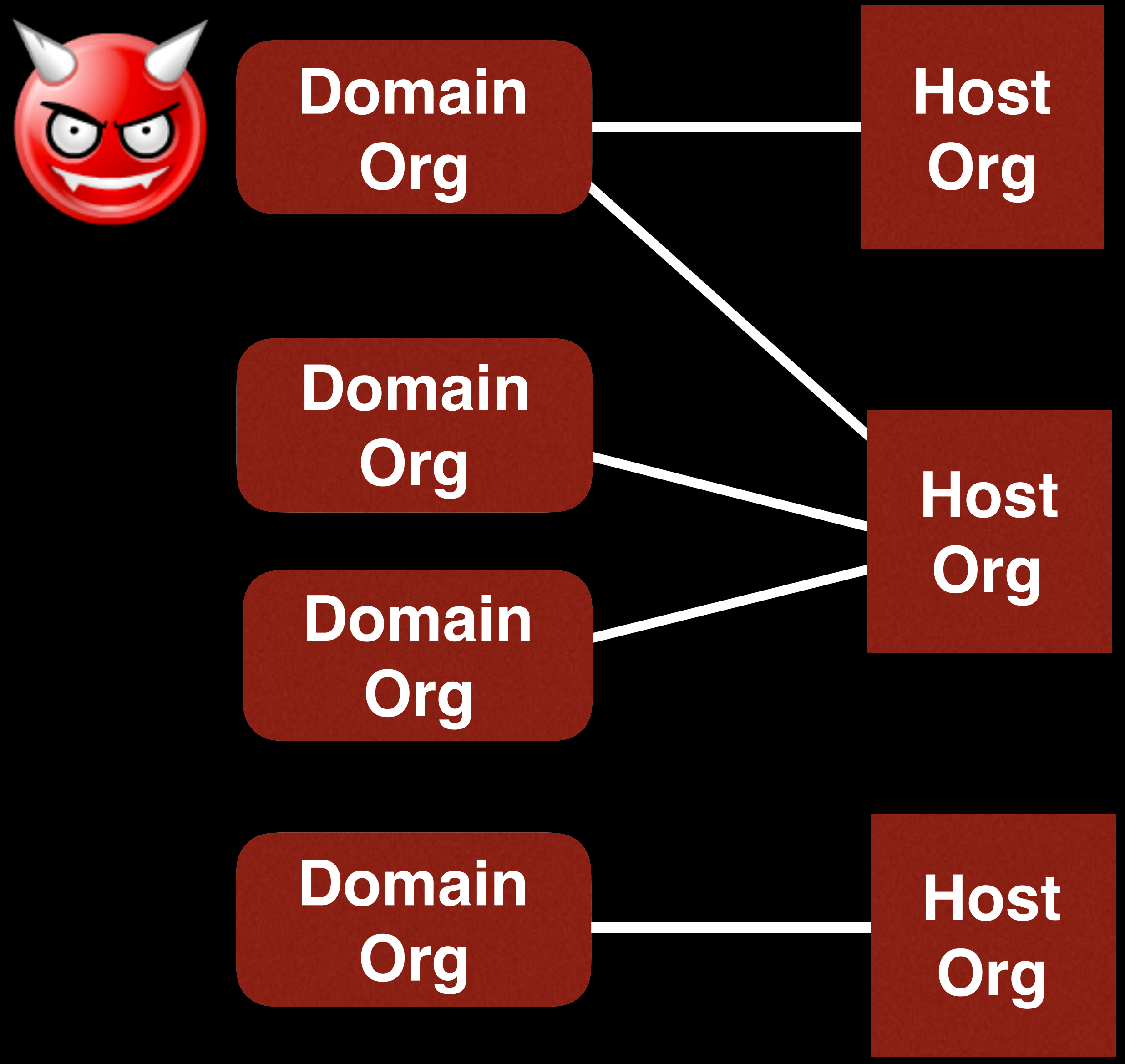
How have keys been aggregated?



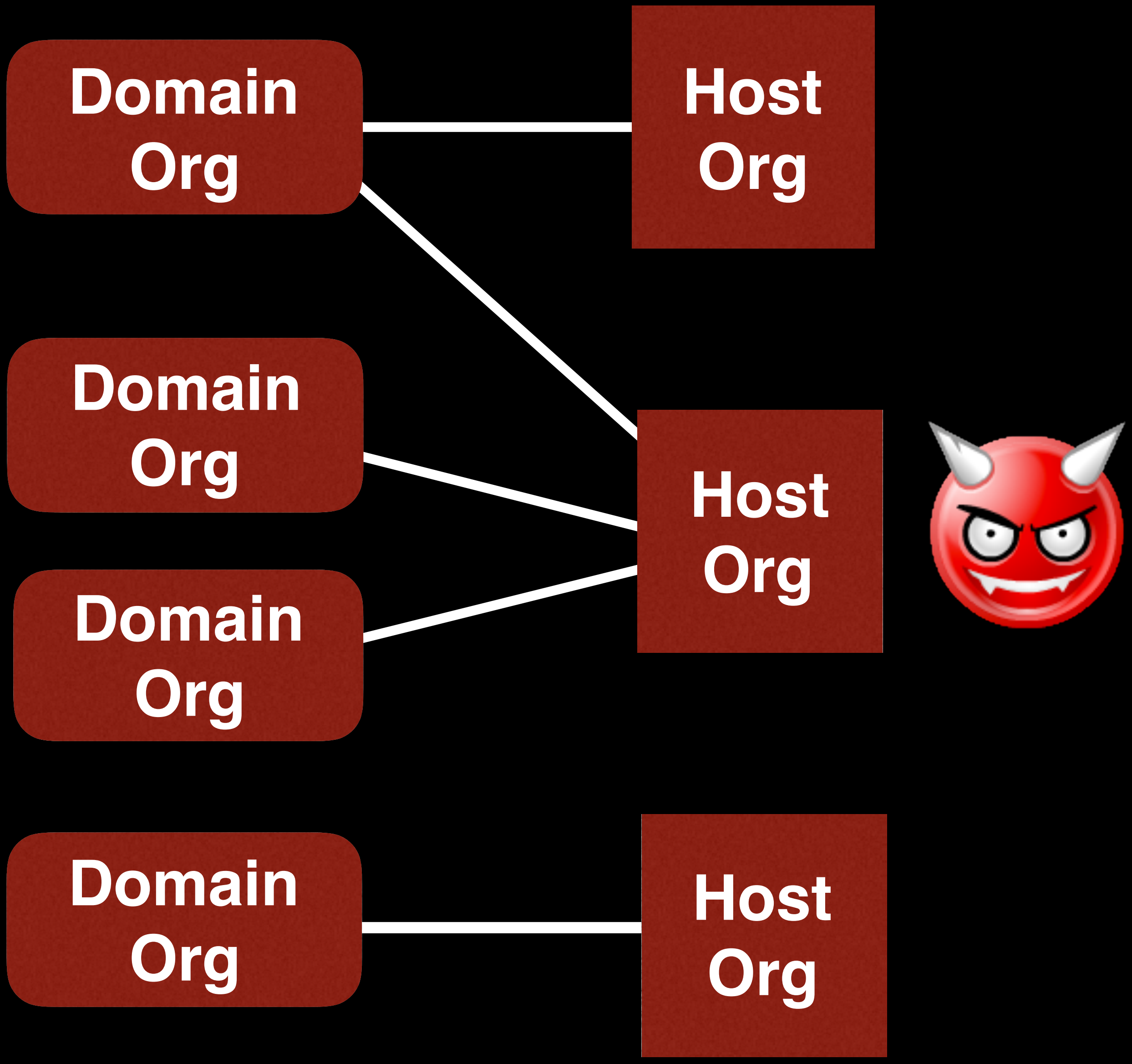
#Organizations	#Domains	Hosting provider
266,110	277,891	secureserver.net
151,628	175,089	amazonaws.com
117.229	122,158	unifiedlayer.com
78,369	87,077	Cloud Flare Inc.
54,158	63,418	Rackspace Hosting
…	…	…
15,440	22,671	akamaitechnologies.com

Top 1% of providers hold keys for 86% of all organizations

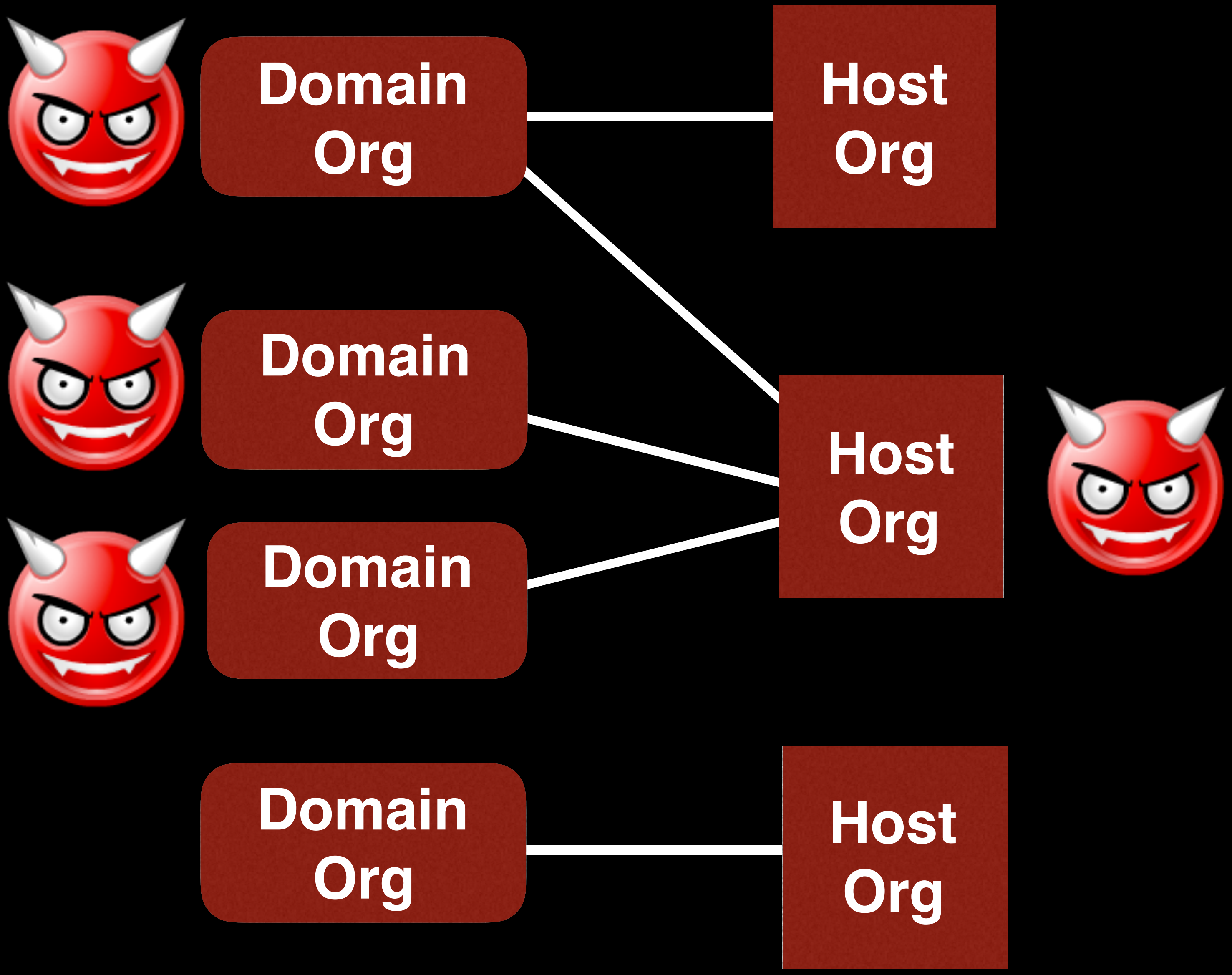
Does key sharing make enticing attack targets?



Does key sharing make enticing attack targets?



Does key sharing make enticing attack targets?



Domain Org

Domain Org

Domain Org

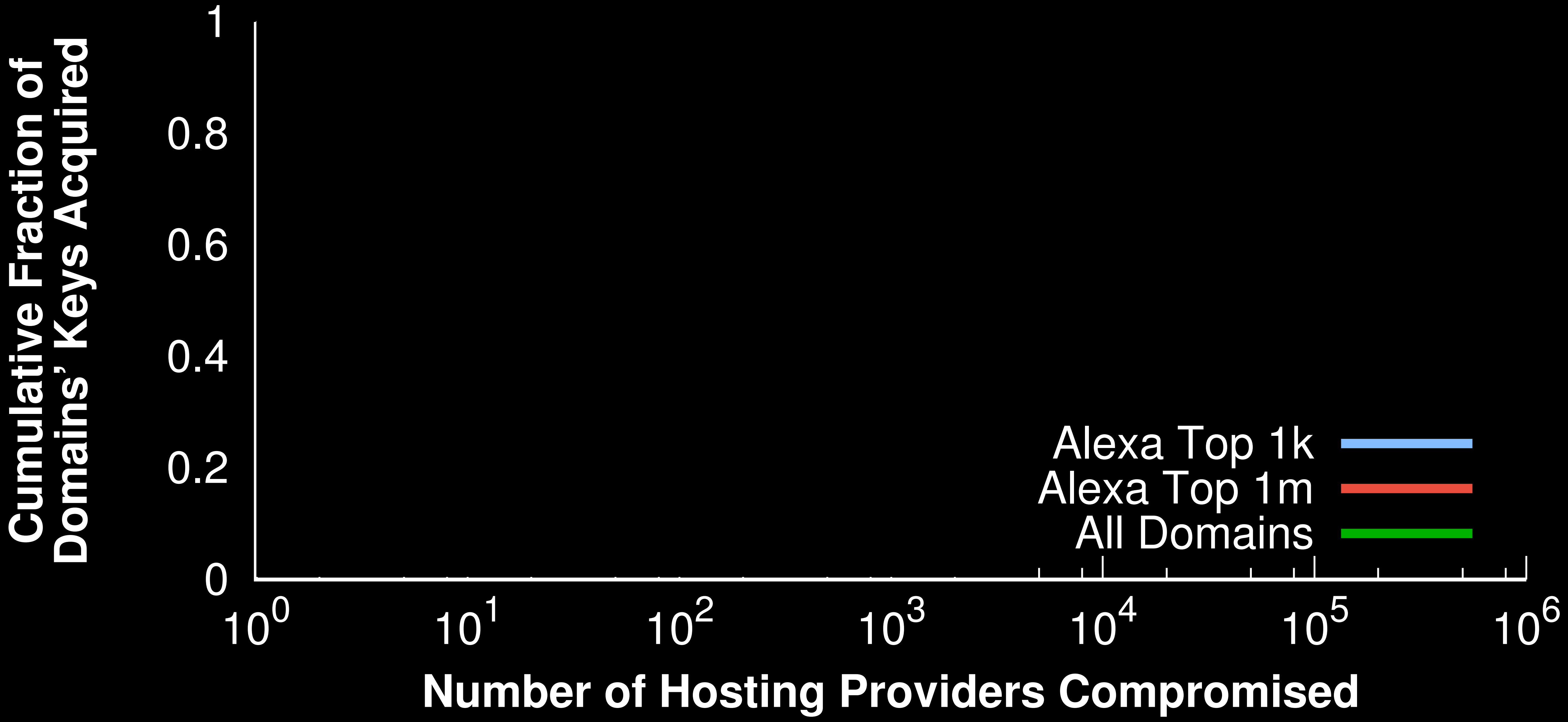
Domain Org

Host Org

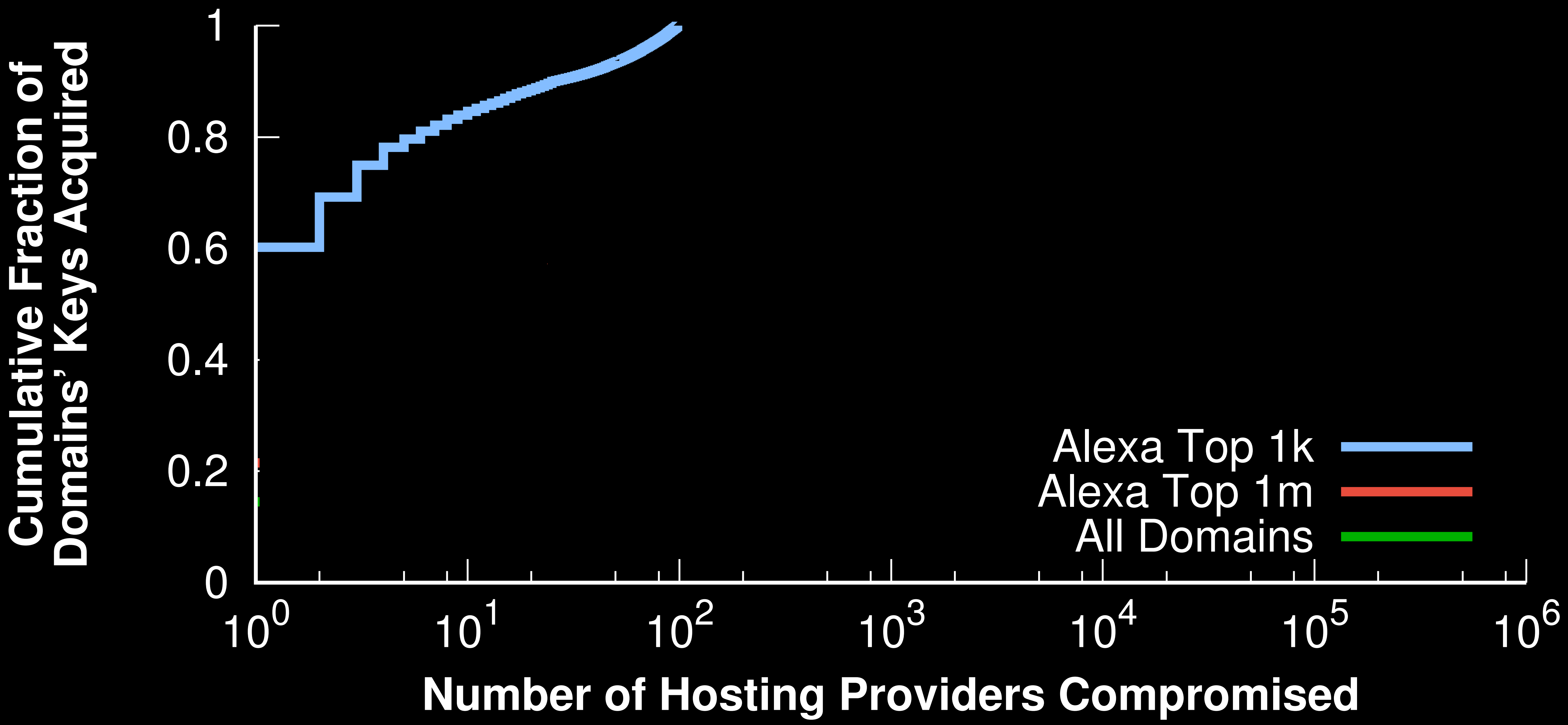
Host Org

Host Org

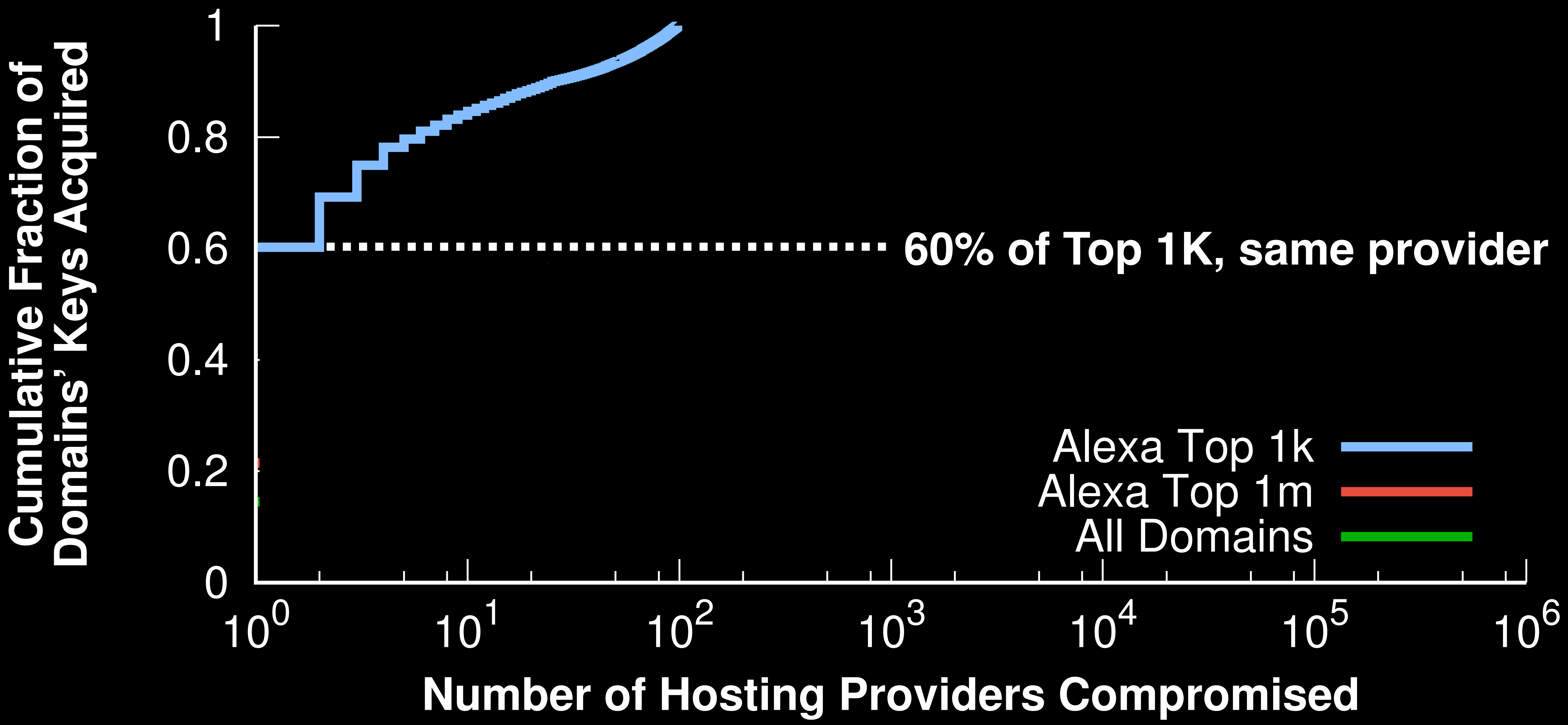
Does key sharing make enticing attack targets?



Does key sharing make enticing attack targets?

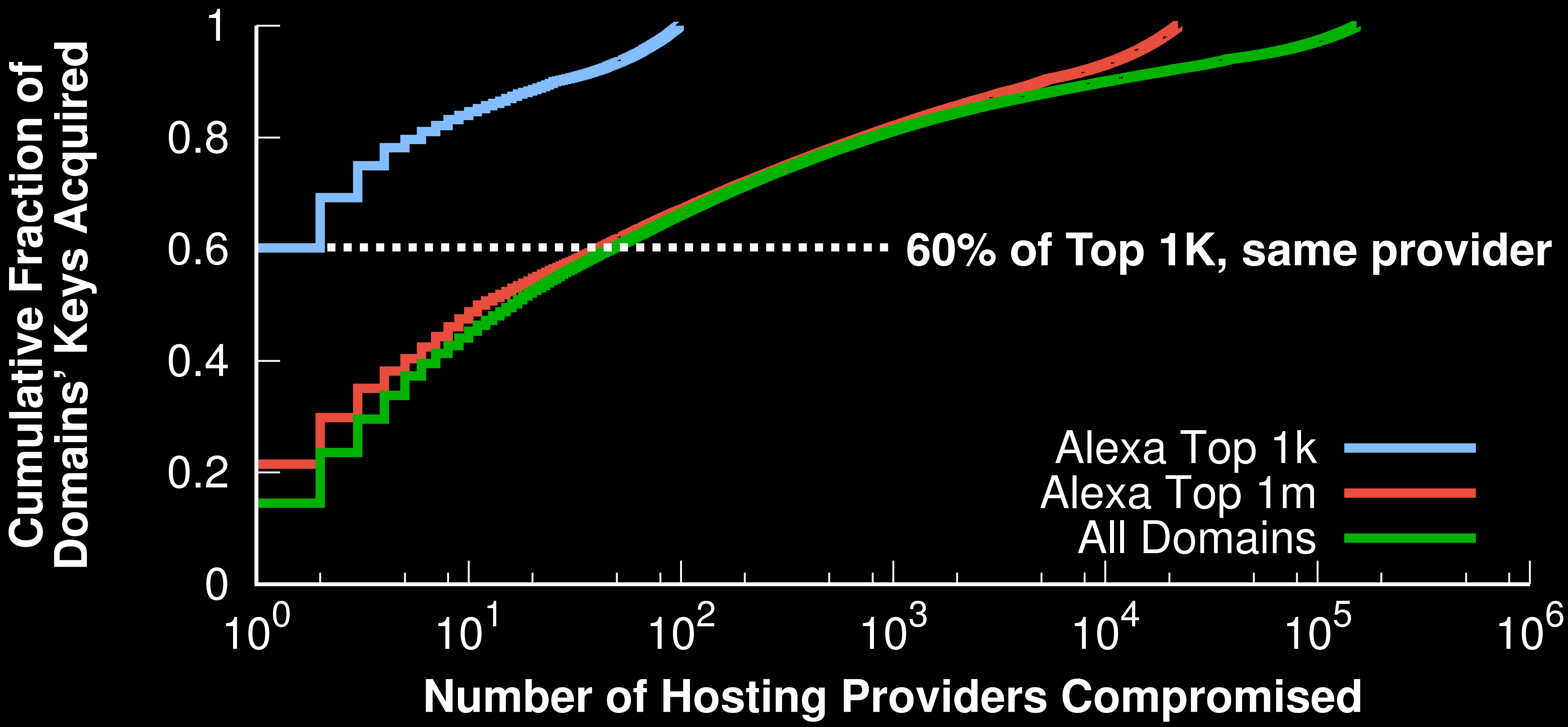


Does key sharing make enticing attack targets?

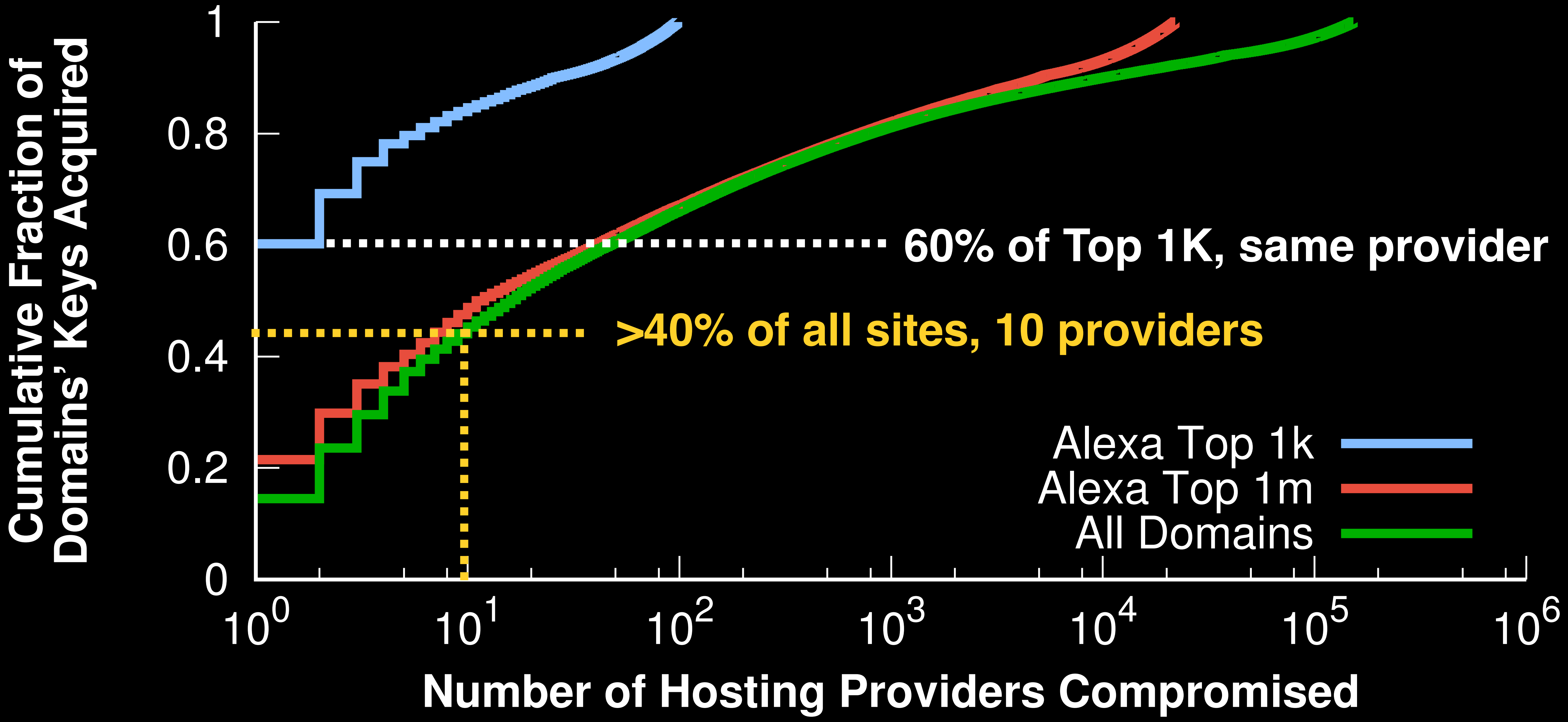


60% of Top 1K, same provider

Does key sharing make enticing attack targets?



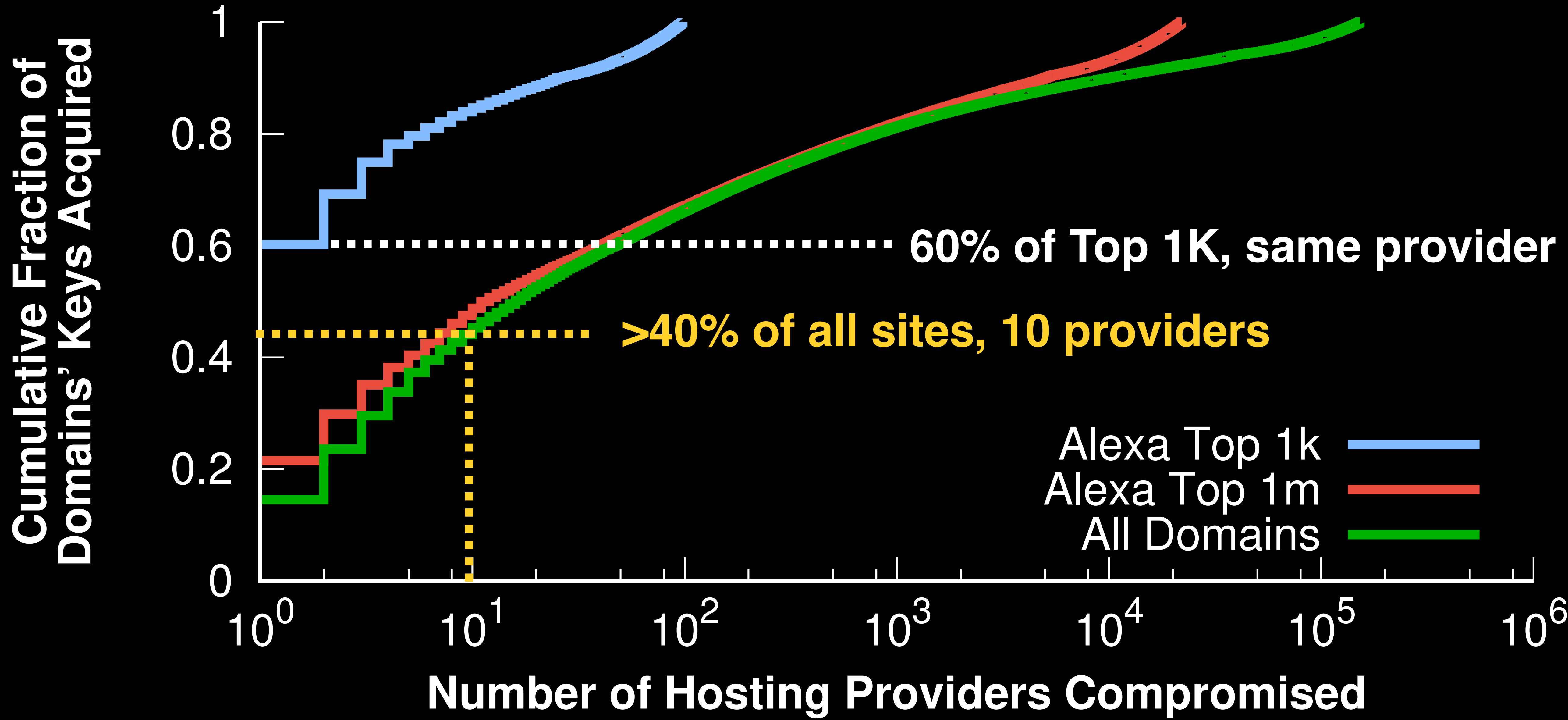
Does key sharing make enticing attack targets?



60% of Top 1K, same provider

>40% of all sites, 10 providers

Does key sharing make enticing attack targets?



Popular hosting services are prime targets for attack

Outline

How prevalent is key sharing?

- 76.5% share with ≥ 1 provider
- Common even among most popular websites

How many keys have providers aggregated?

- Top 1% of providers hold keys for 86% of orgs
- Attractive targets for attack

How does sharing impact key management?

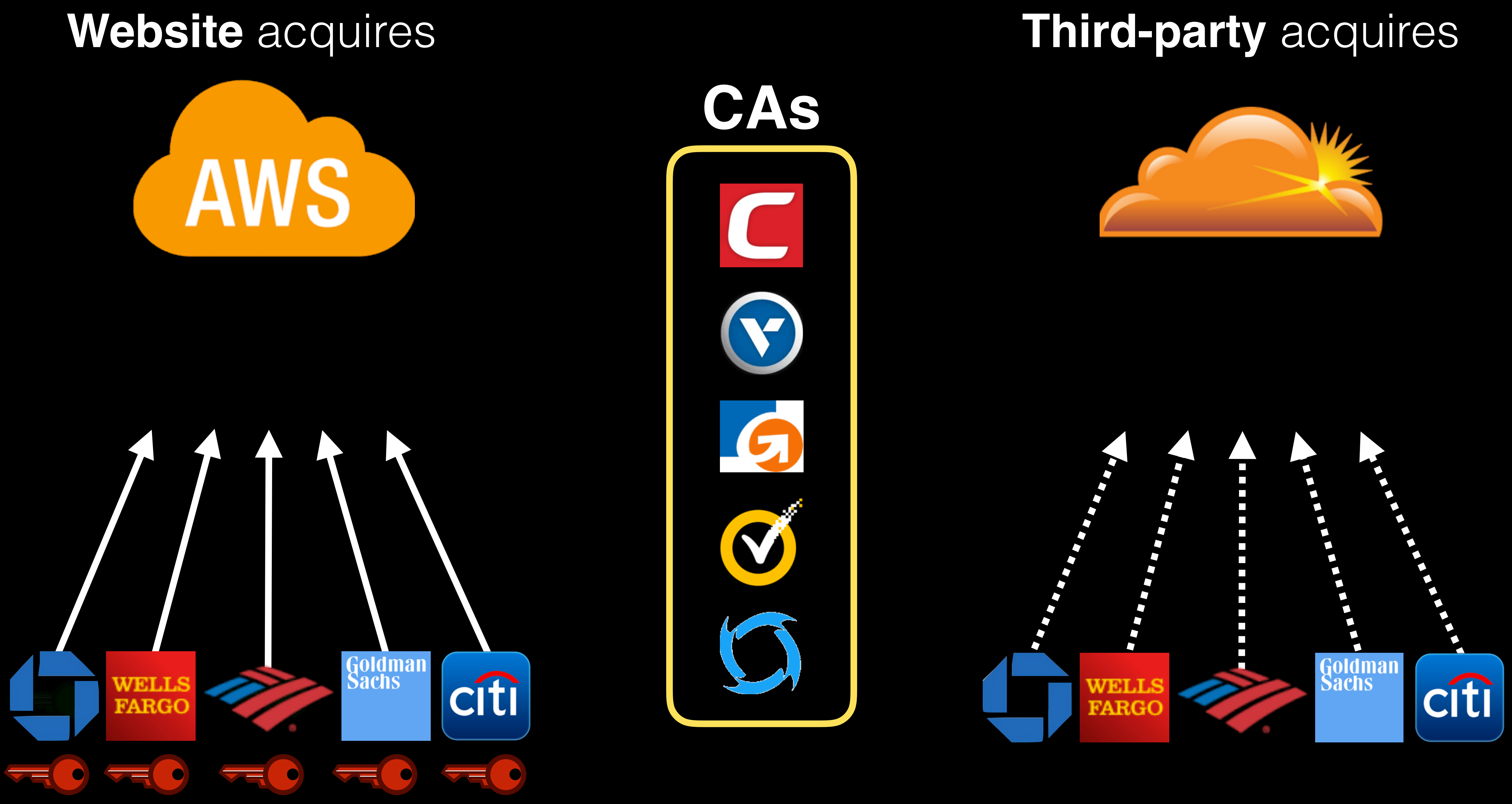
Key Management

Request certificates

Renew expiring certificates

Revoke and reissue compromised certificates

Who manages private keys?



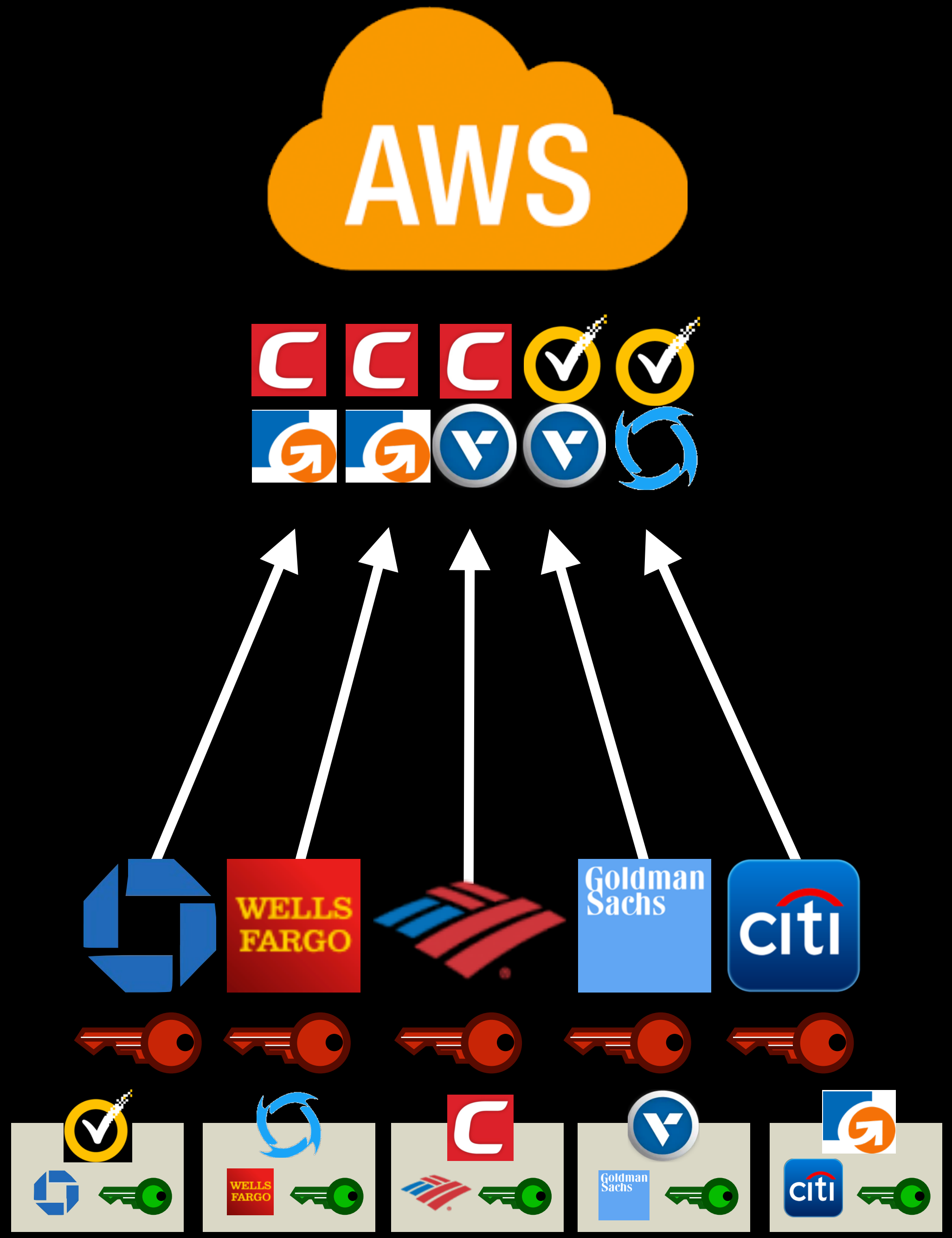
Website acquires

CAs

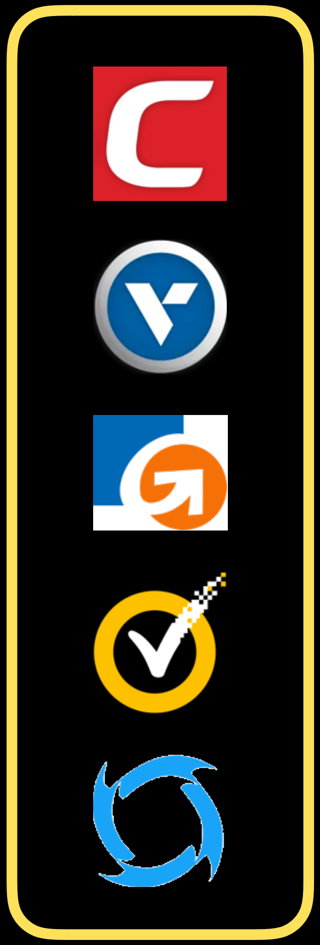
Third-party acquires

Who manages private keys?

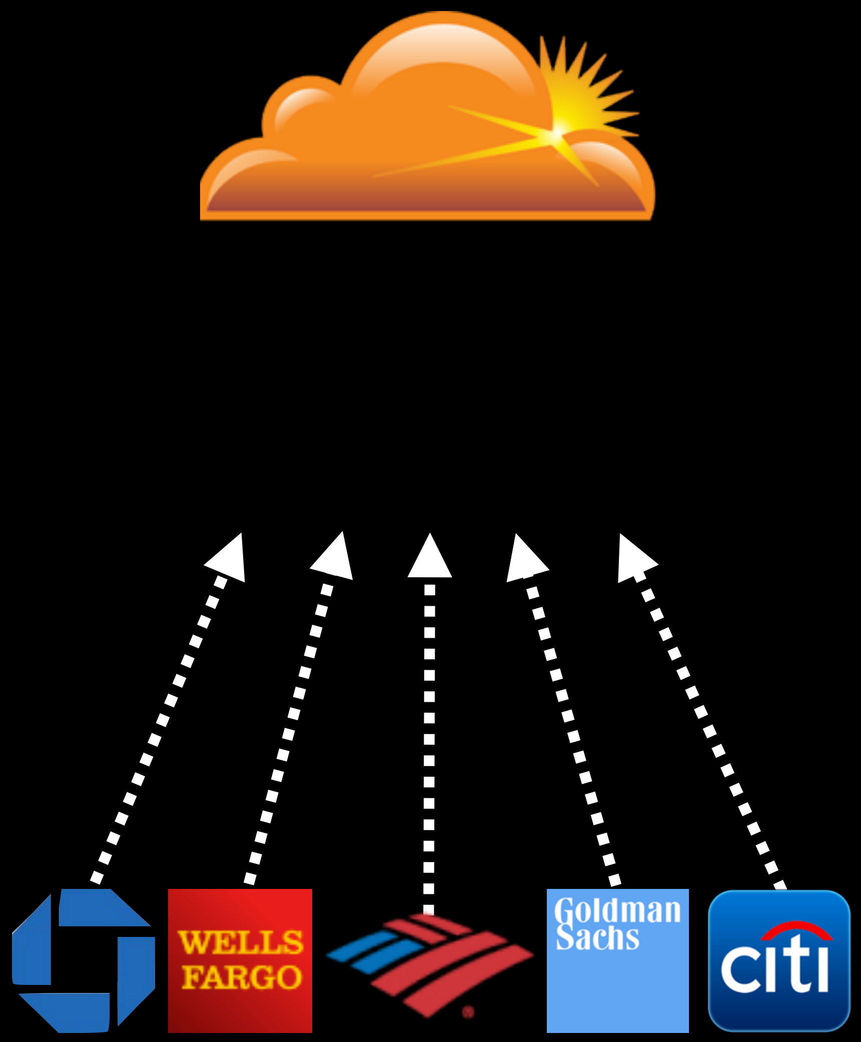
Website acquires



CAs



Third-party acquires

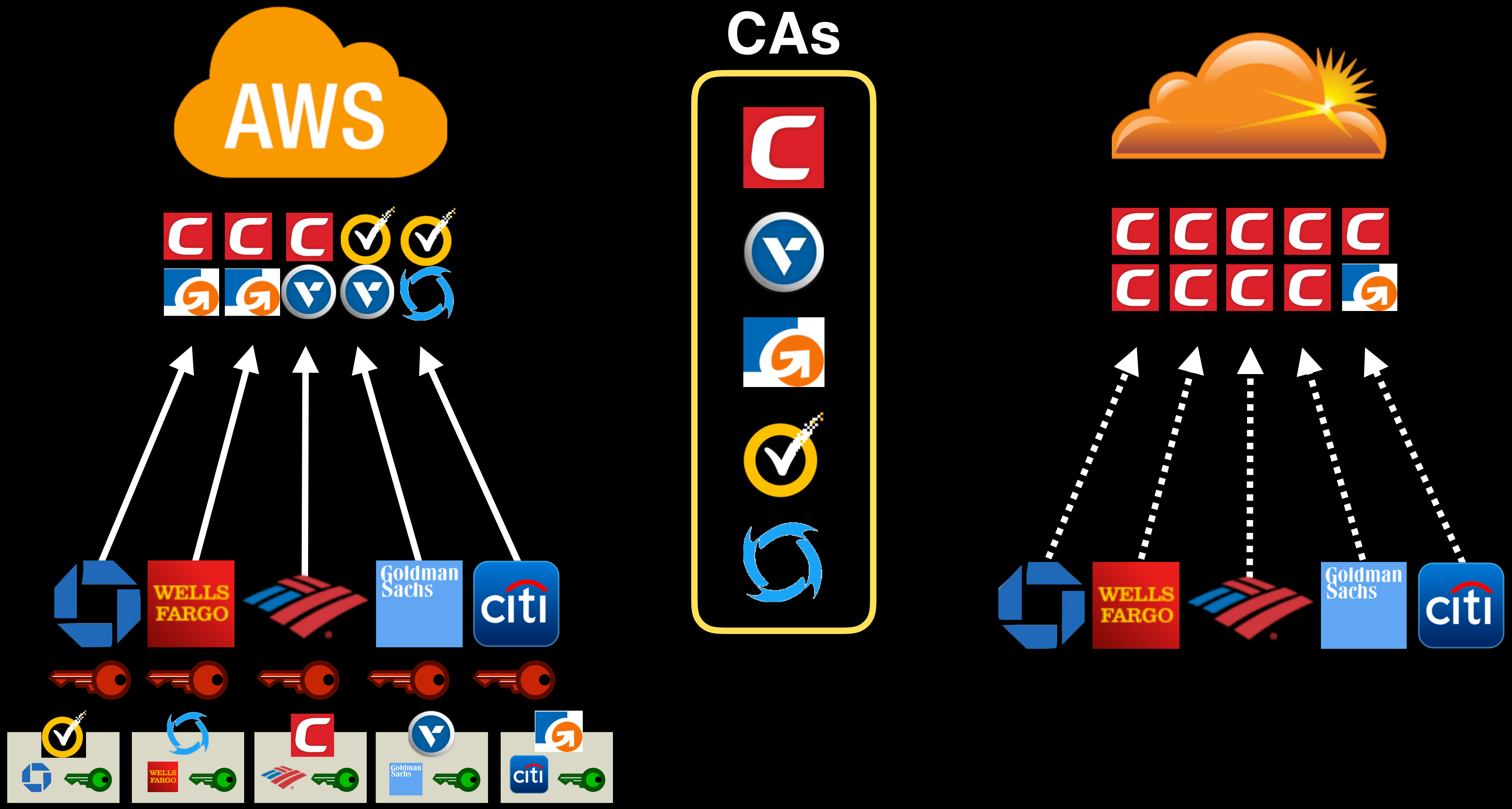


Who manages private keys?

Website acquires

CAs

Third-party acquires



Who manages private keys?

Website acquires

Third-party acquires

Who manages private keys?

Website acquires

Diverse

↓

“Self-managed”

Third-party acquires

Who manages private keys?

Website acquires

Diverse

↓

“Self-managed”

Third-party acquires

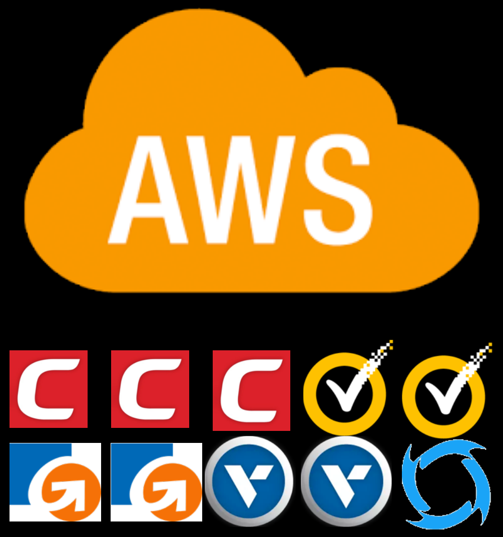
Heavily skewed

↓

“Outsourced”

Who manages private keys?

Website acquires

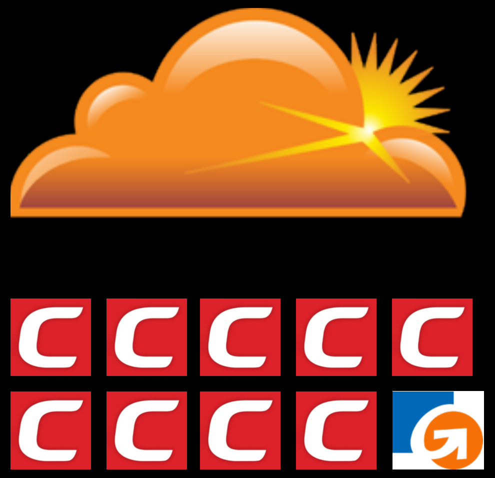


Diverse

↓

“Self-managed”

Third-party acquires



Heavily skewed

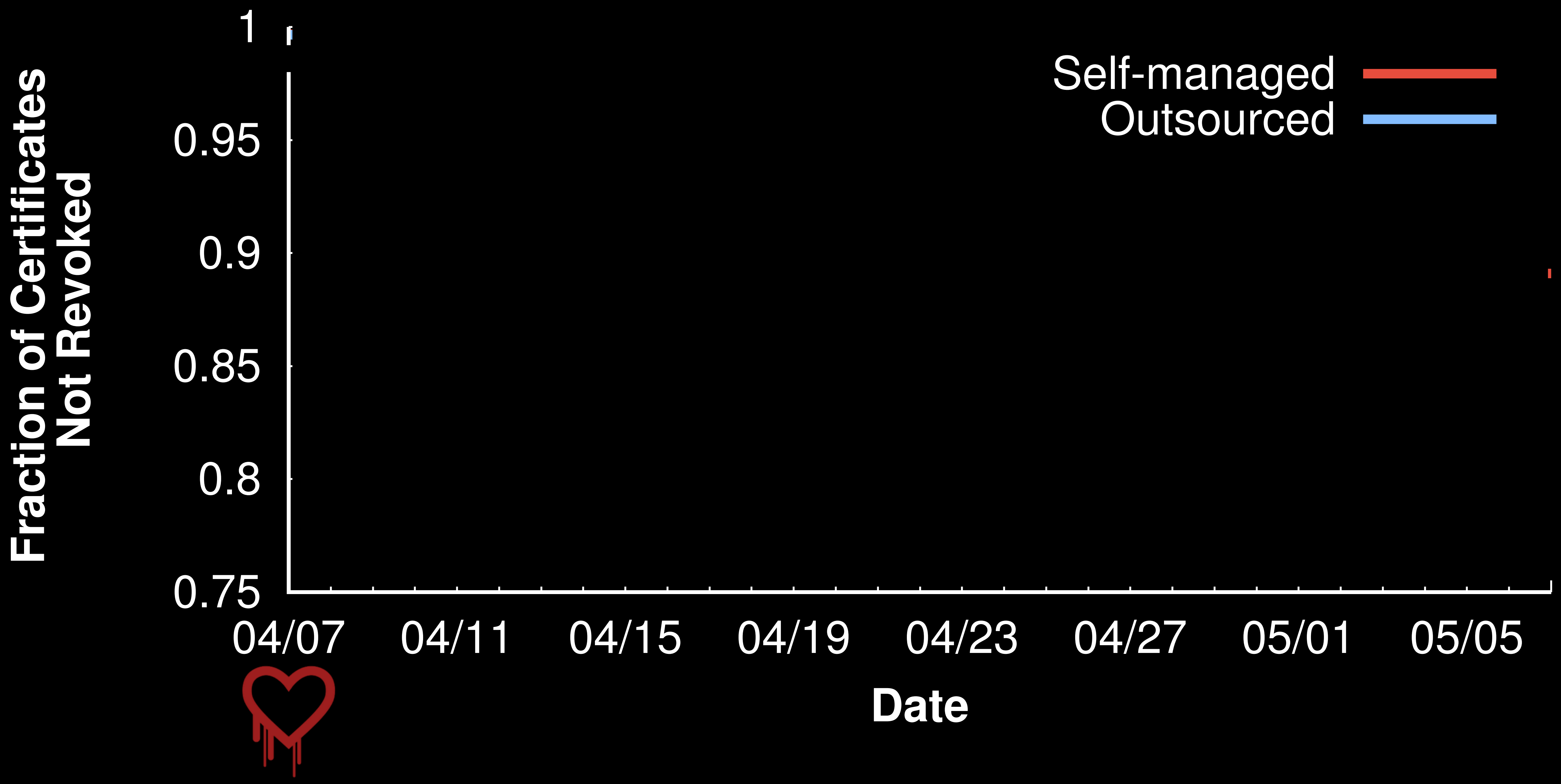
↓

“Outsourced”

58.4% of Alexa Top 10K
33.0% of all domains

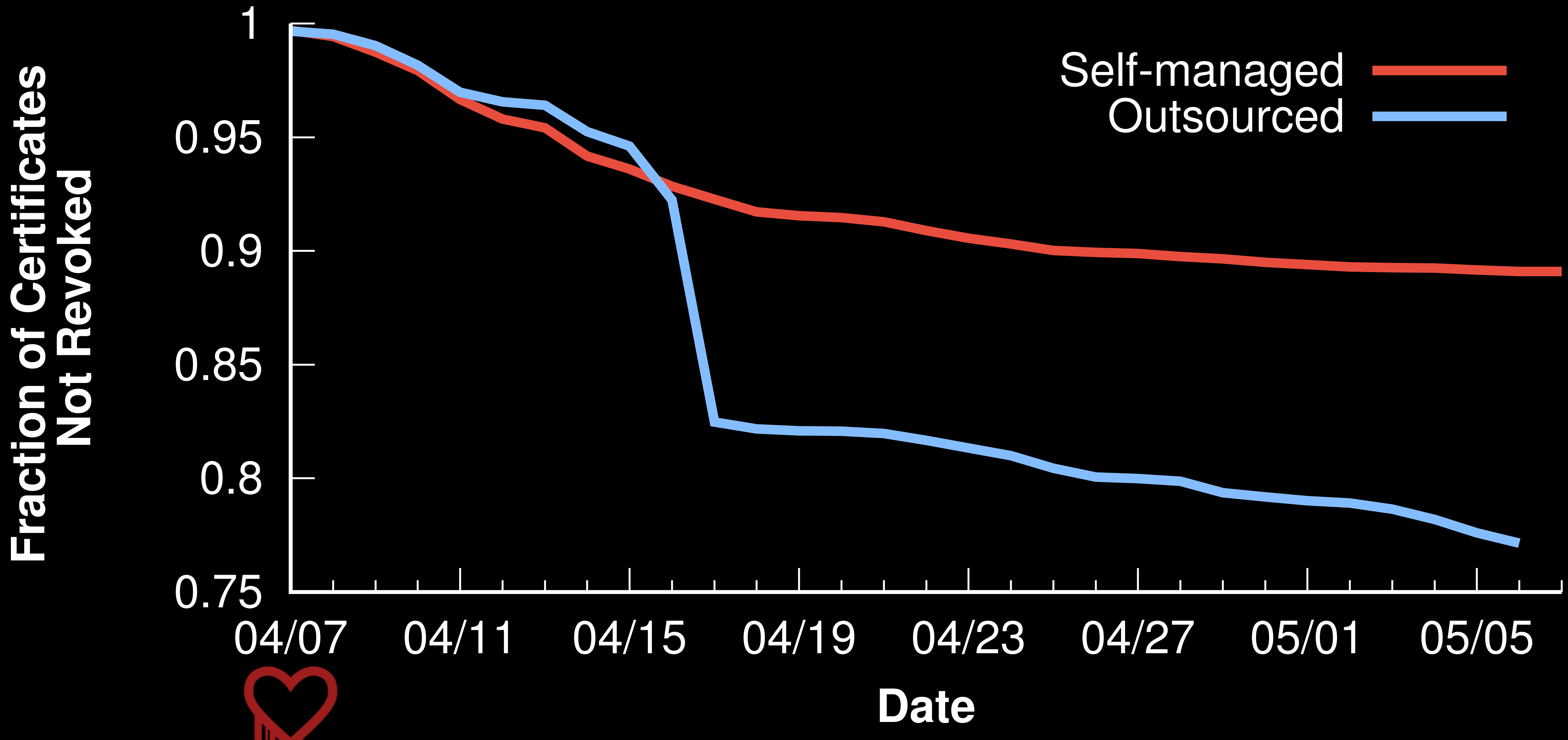
How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)



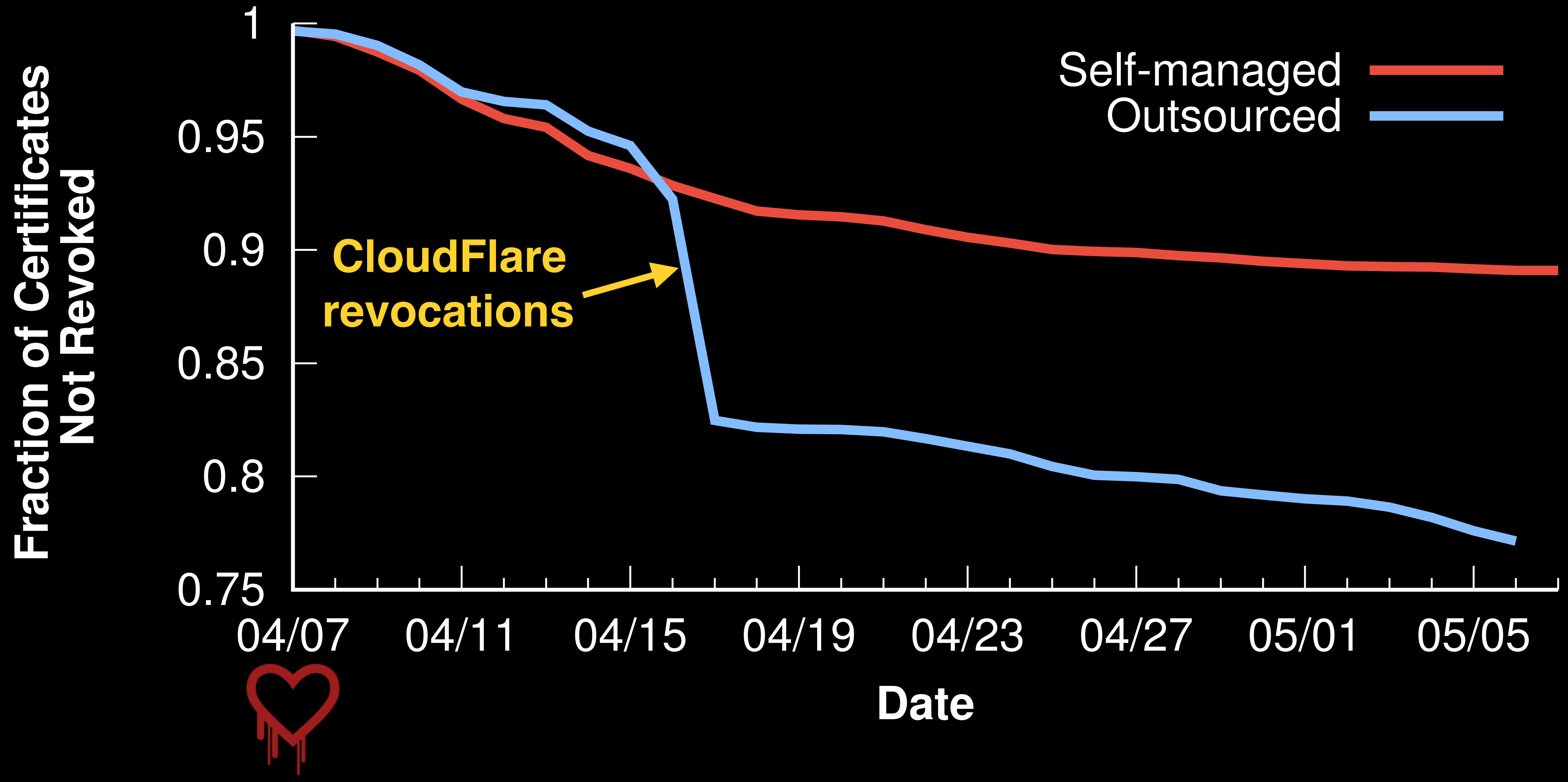
How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)



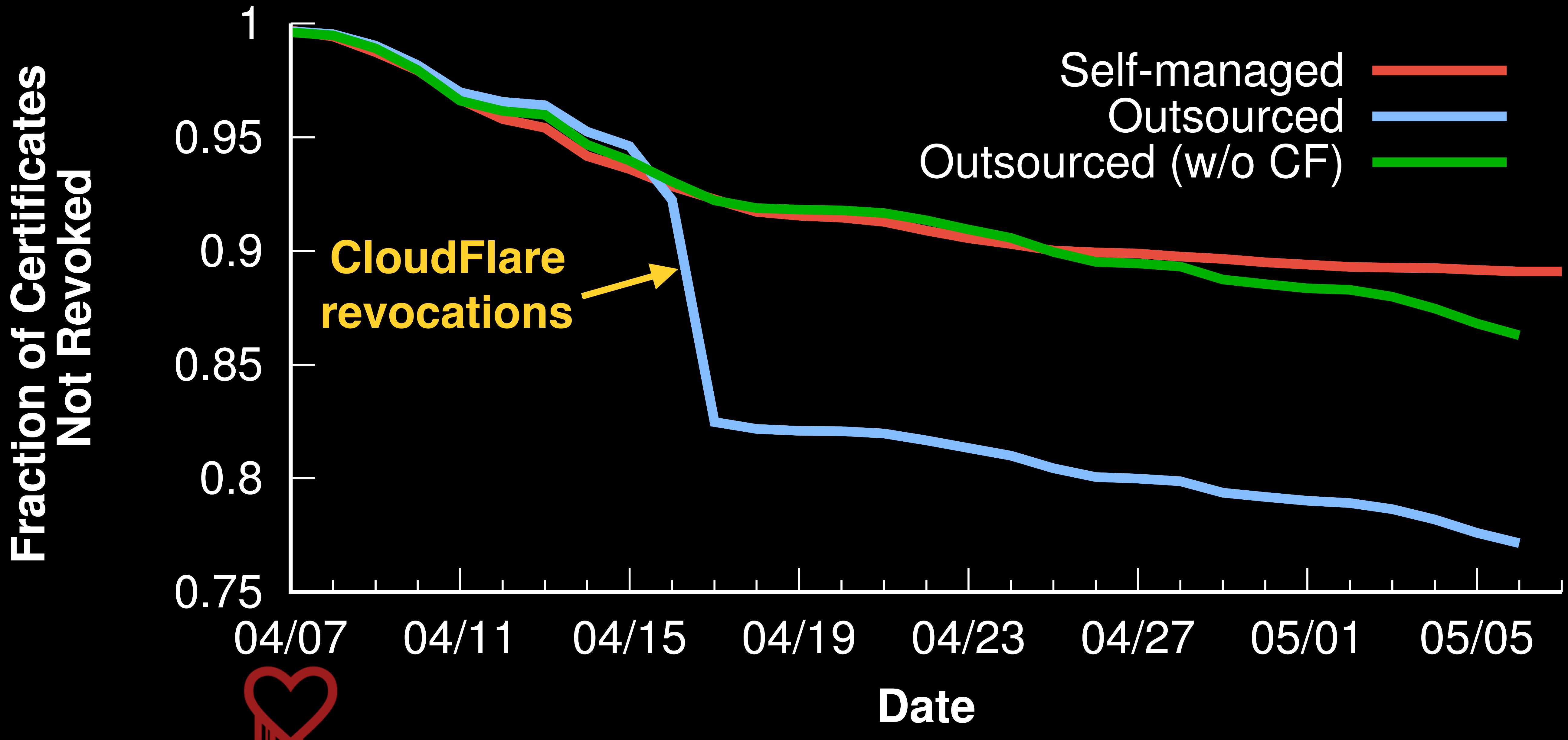
How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)



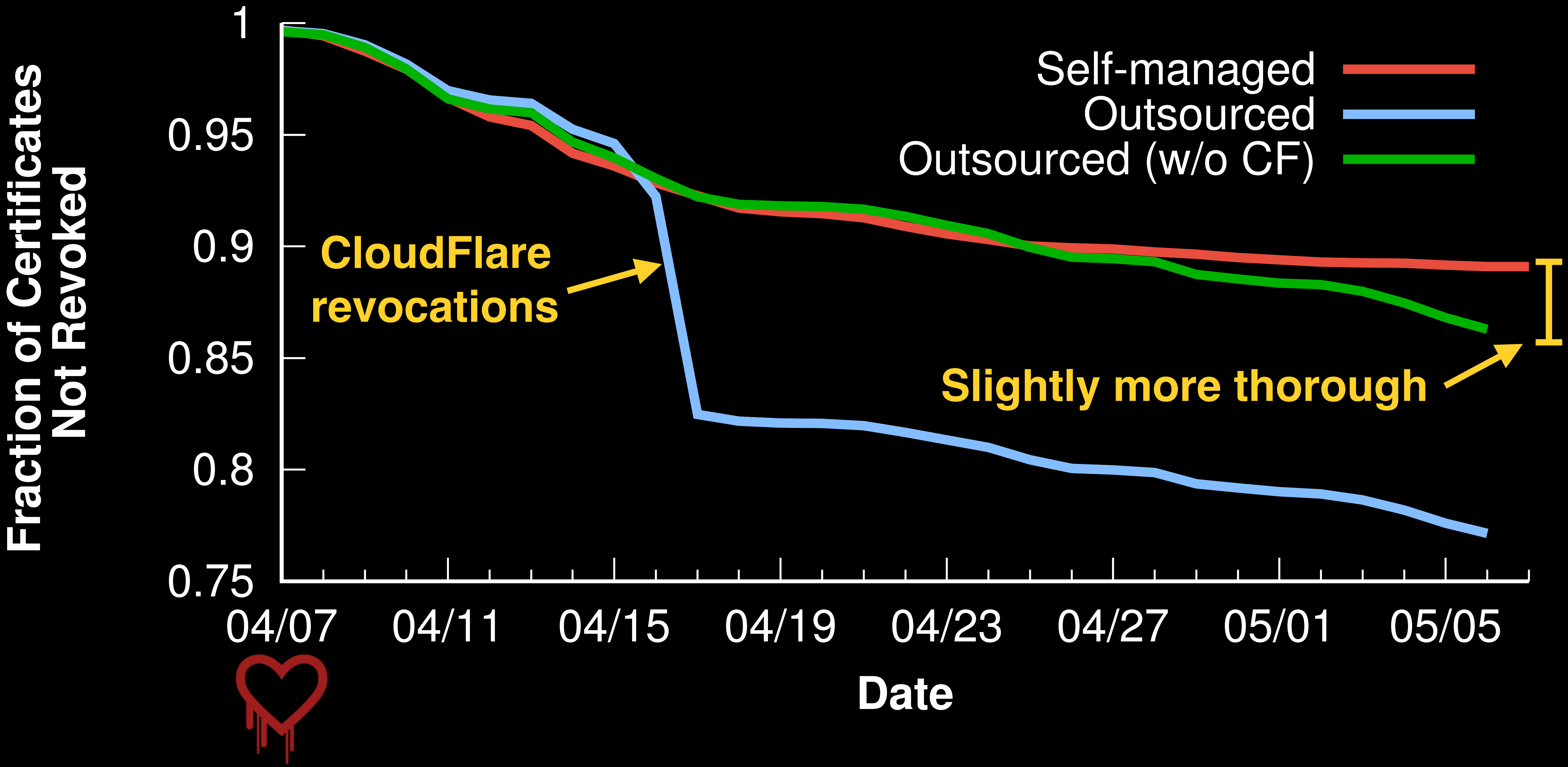
How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)



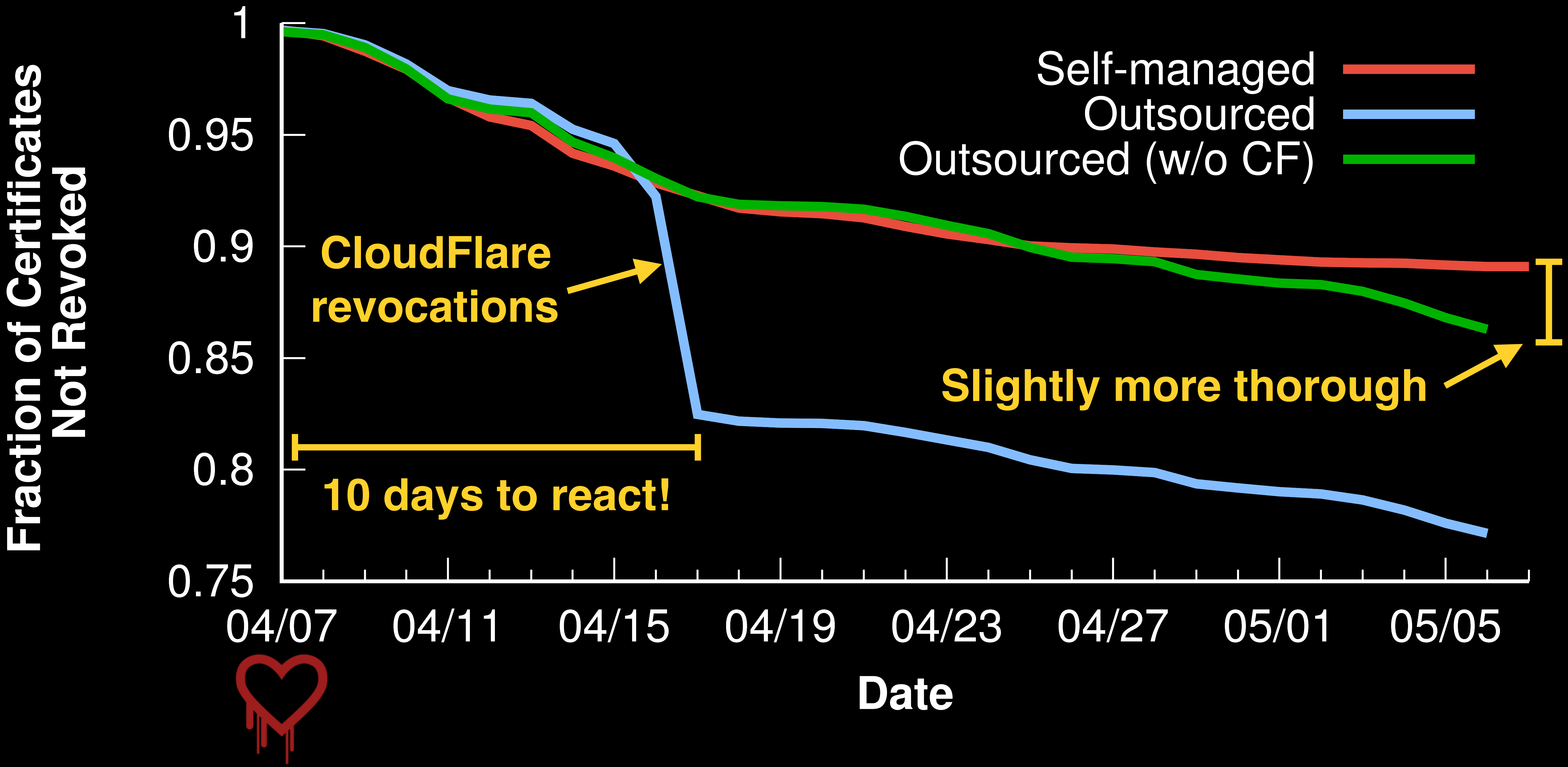
How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)



How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)



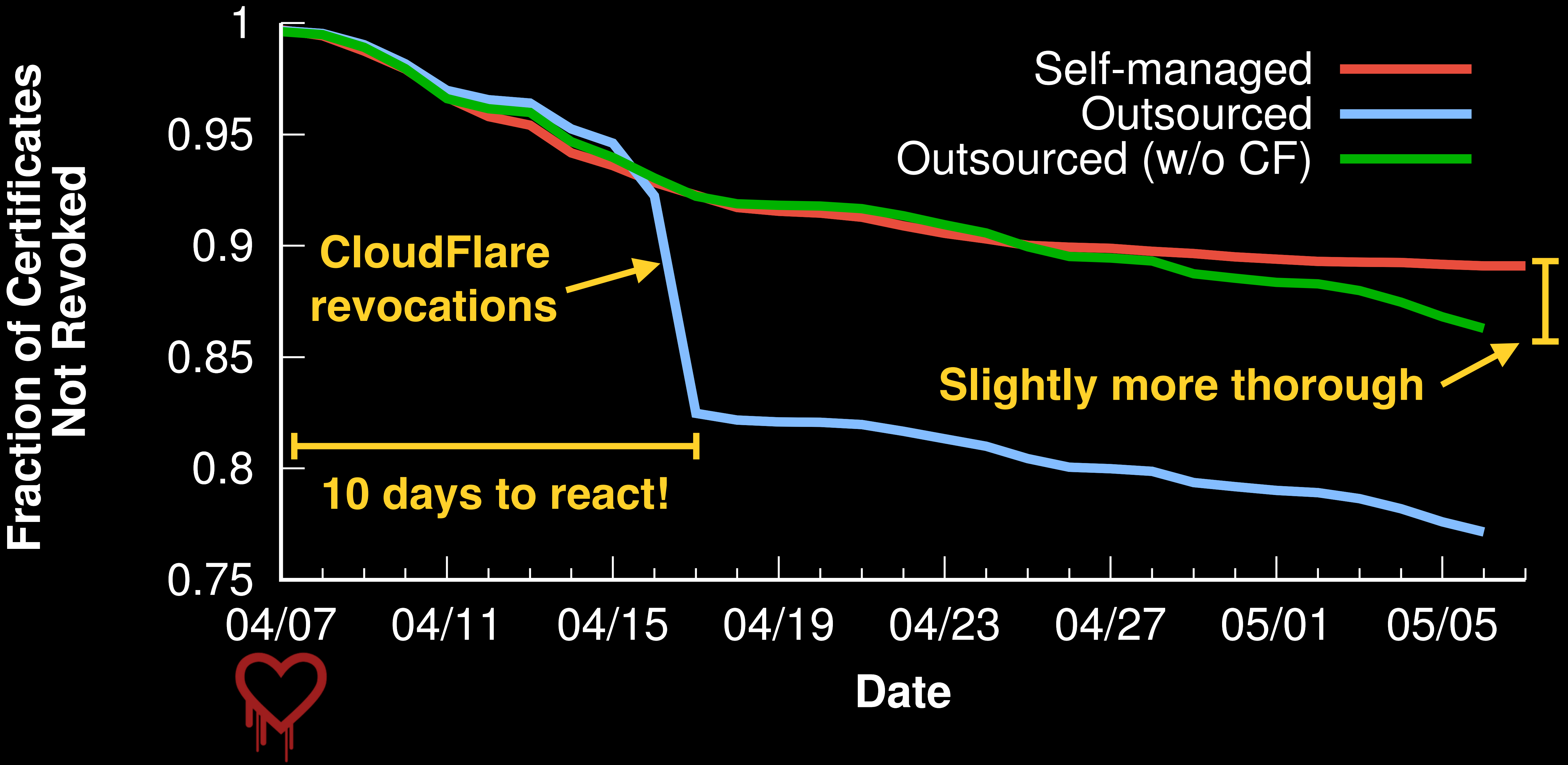
CloudFlare revocations

10 days to react!

Slightly more thorough

How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)



A few revoked thoroughly, but many did not!

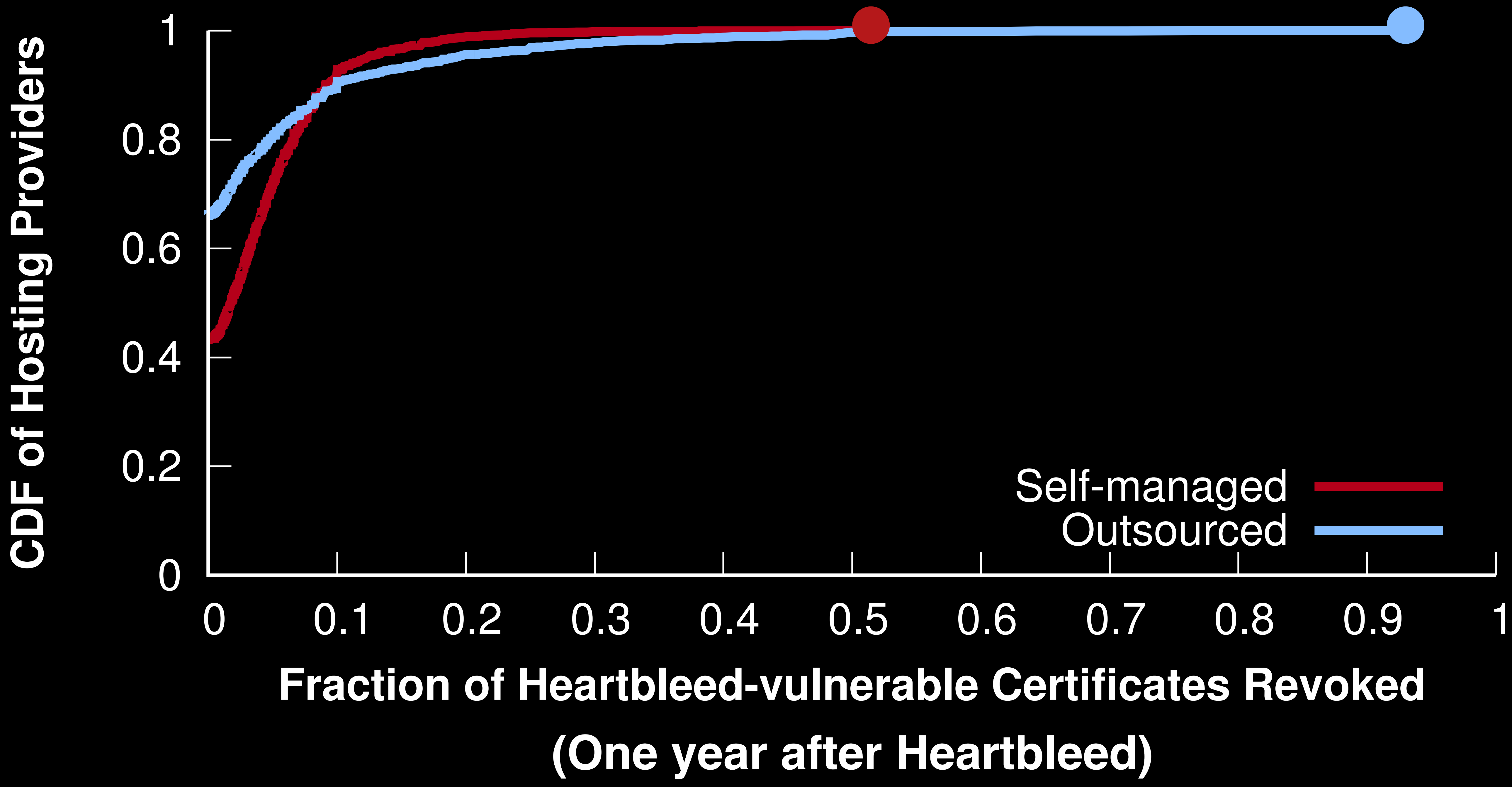
How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)

CDF of Hosting Providers
1
0.8
0.6
0.4
0.2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Self-managed
Outsourced
Fraction of Heartbleed-vulnerable Certificates Revoked
(One year after Heartbleed)

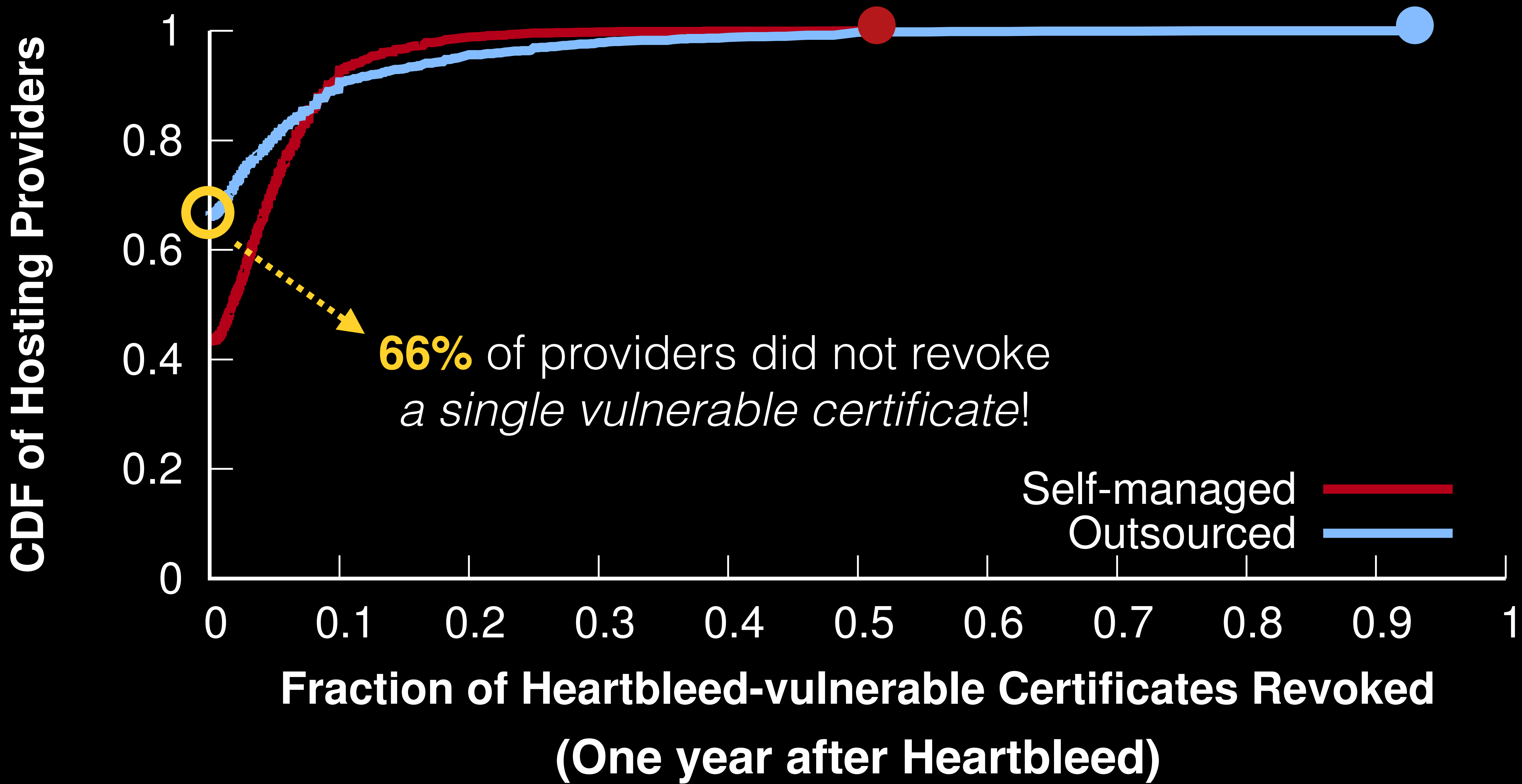
How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)



How does sharing impact key management?

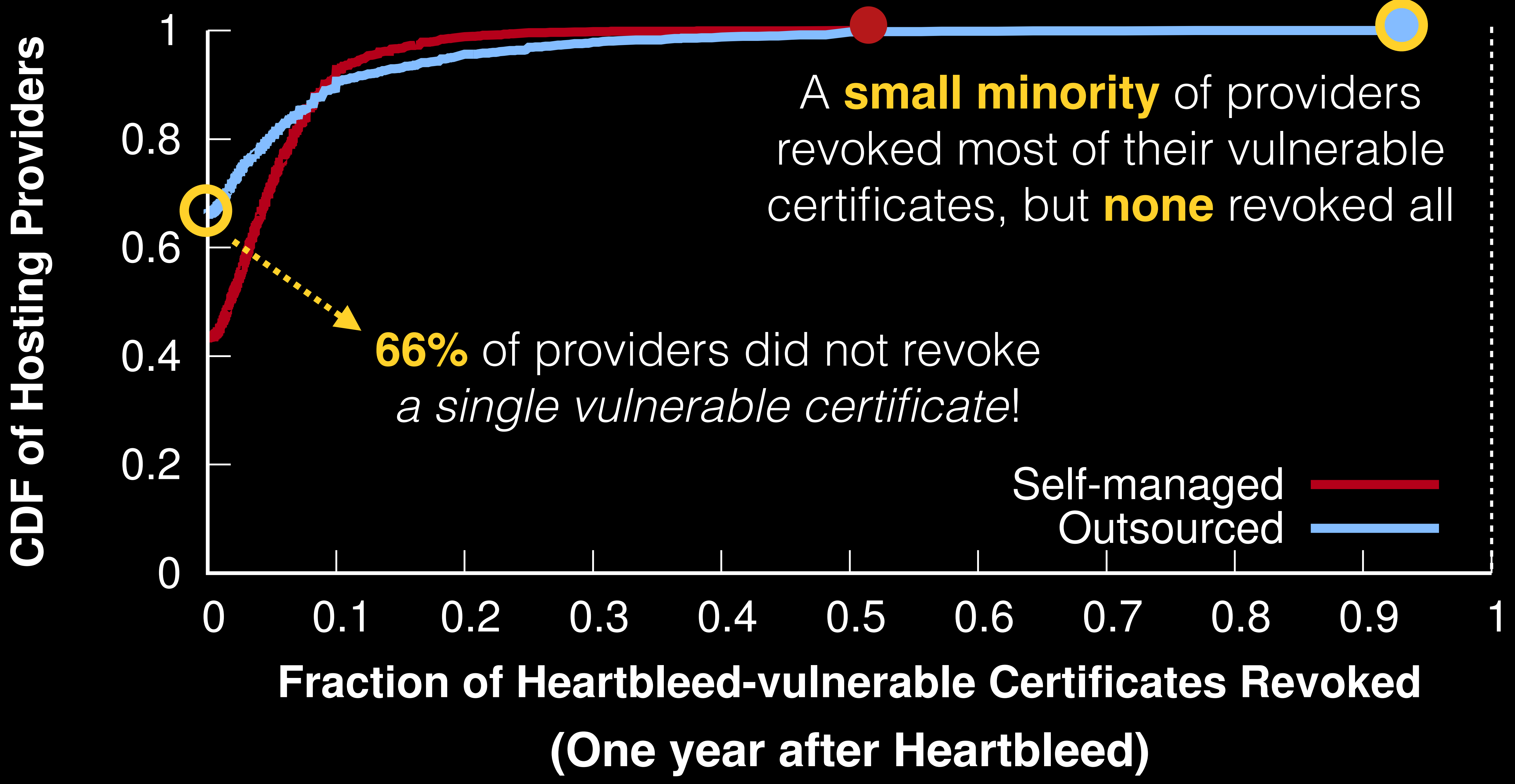
Natural experiment: Heartbleed (4/7/2014)



66% of providers did not revoke *a single vulnerable certificate*!

How does sharing impact key management?

Natural experiment: Heartbleed (4/7/2014)



A **small minority** of providers revoked most of their vulnerable certificates, but **none** revoked all

66% of providers did not revoke *a single vulnerable certificate*!

Outline

How prevalent is key sharing?

- 76.5% share with ≥ 1 provider
- Common even among most popular websites

How many keys have providers aggregated?

- Top 1% of providers hold keys for 86% of orgs
- Attractive targets for attack

How does sharing impact key management?

- Creates single point of failure
- Most third-parties did poor job of revoking

Due to **economic incentives**,
key sharing is prevalent in today's web

Most providers are *not* managing keys responsibly

Future work on the PKI should take *economics* and hosting providers into account, ideally: hosting should not *require* key sharing

frankc@csail.mit.edu

Code and data available at:

securepki.org