

Who's Got Your Mail?

Characterizing Mail Service Provider Usage

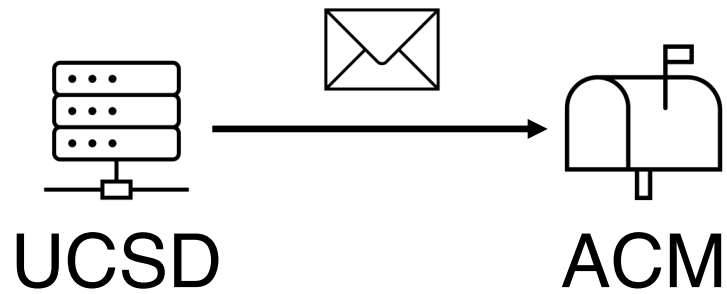
Enze “Alex” Liu, Gautam Akiwate, Mattijs Jonker[†], Ariana Mirian,
Stefan Savage, Geoffrey M. Voelker

UC San Diego

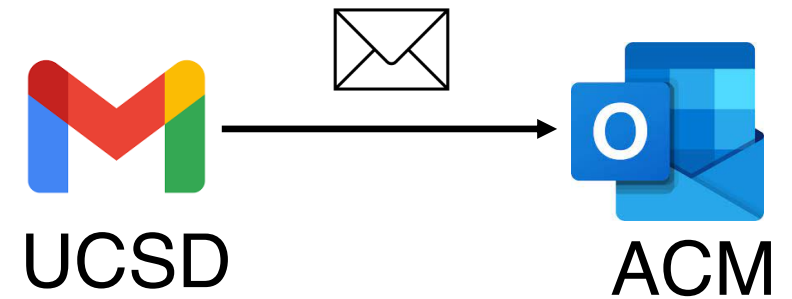
[†]University of Twente

Mail Service Centralization

Traditionally



Now



Mail Service Centralization

- Availability
- Security
- Legal implications

Gmail, YouTube, Google Docs, and other Google services hit by massive outage

Google says the problem should be fixed for most users now

By [Jon Porter](#) | [@JonPorty](#) | Updated Dec 14, 2020, 8:14am EST

2021 Microsoft Exchange Server data breach

From Wikipedia, the free encyclopedia

A global wave of [cyberattacks](#) and [data breaches](#) began in January 2021 after four [zero-day exploits](#) were discovered on affected servers, [administrator privileges](#) on the server, and access to connected devices on the same network.

Stored Communications Act

From Wikipedia, the free encyclopedia

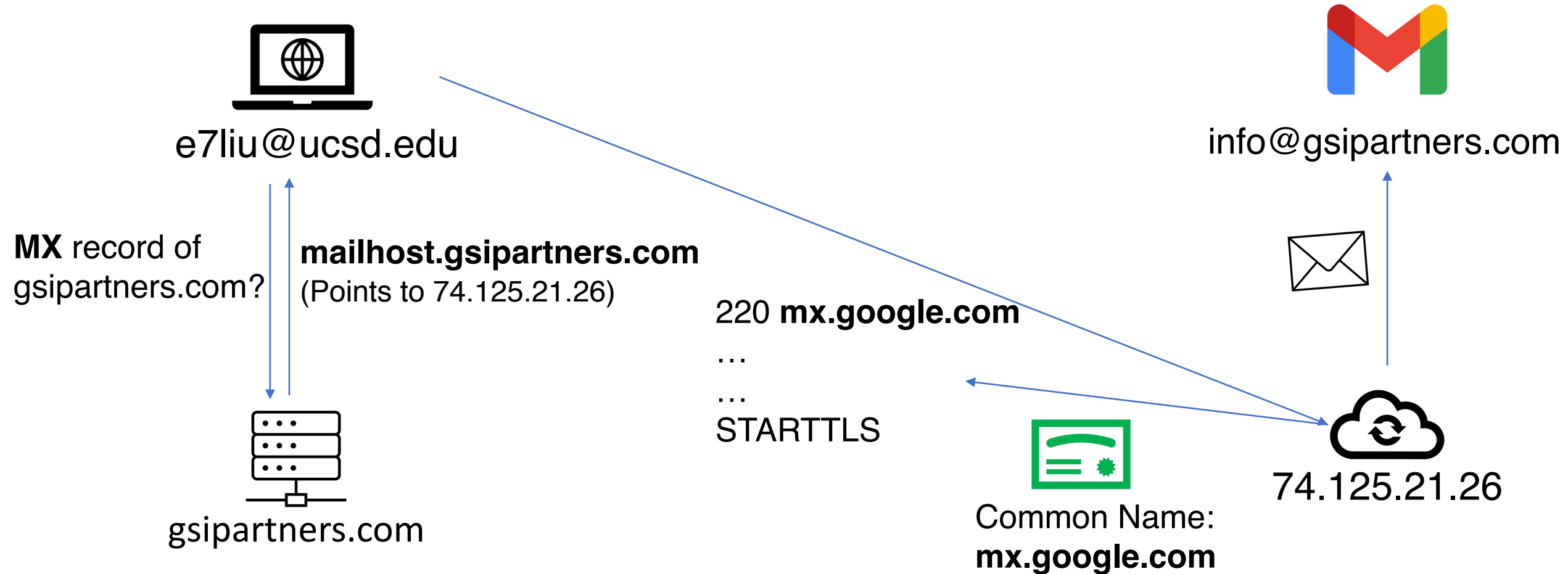
The **Stored Communications Act** (SCA, [codified at 18 U.S.C. Chapter 121](#) §§ 2701–2712) is a federal law that governs the storage of electronic communications by third-party [internet service providers](#) (ISPs). It was enacted as Title 18, Chapter 121 of the United States Code.

Contributions

- **Methodology** for mapping domains to mail service providers
- **Characterization** of mail service provider usage
 - Current landscape
 - Longitudinal trends
 - Provider preference by country

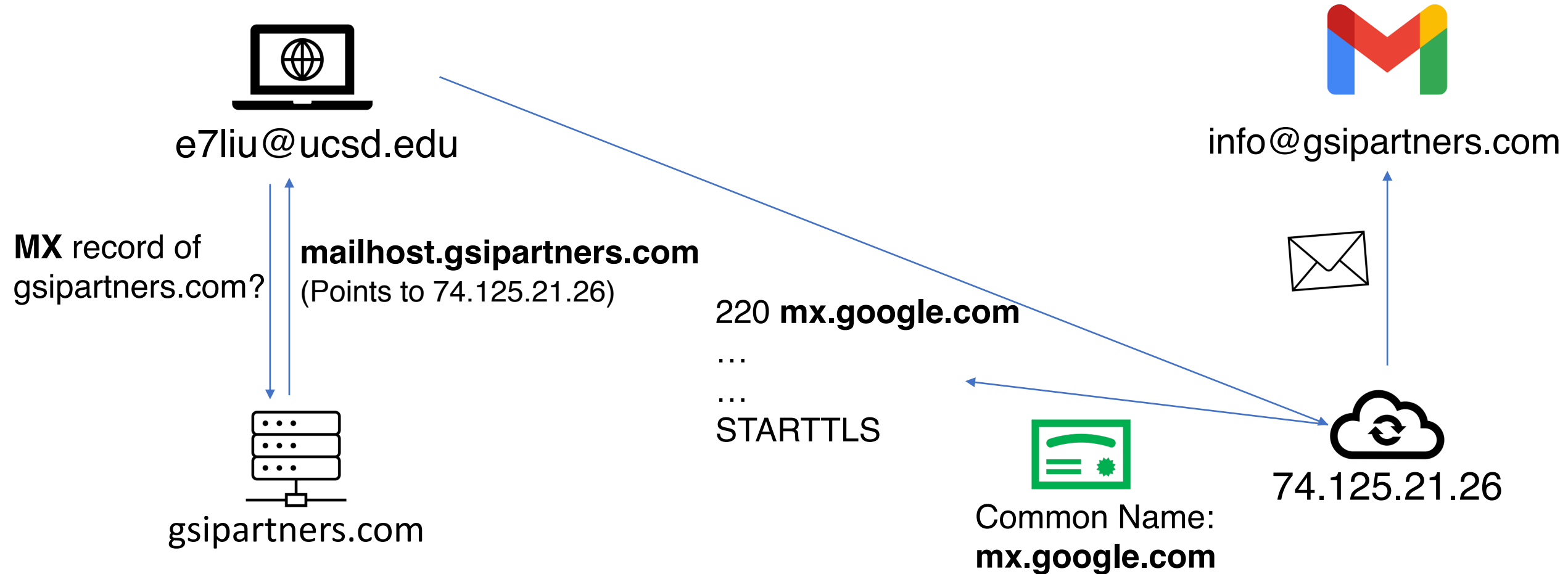
Mail Delivery in a Typical Case

Video Here



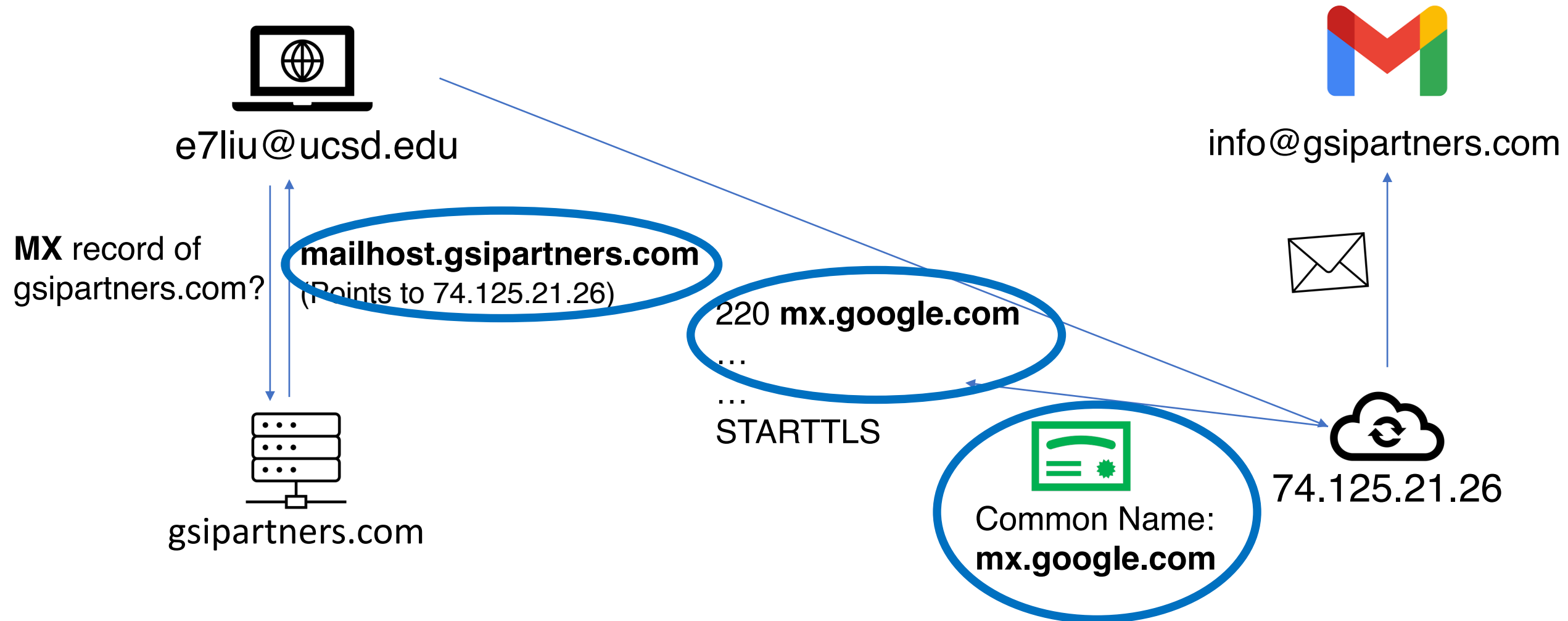
Mail Delivery in a Typical Case

Video Here



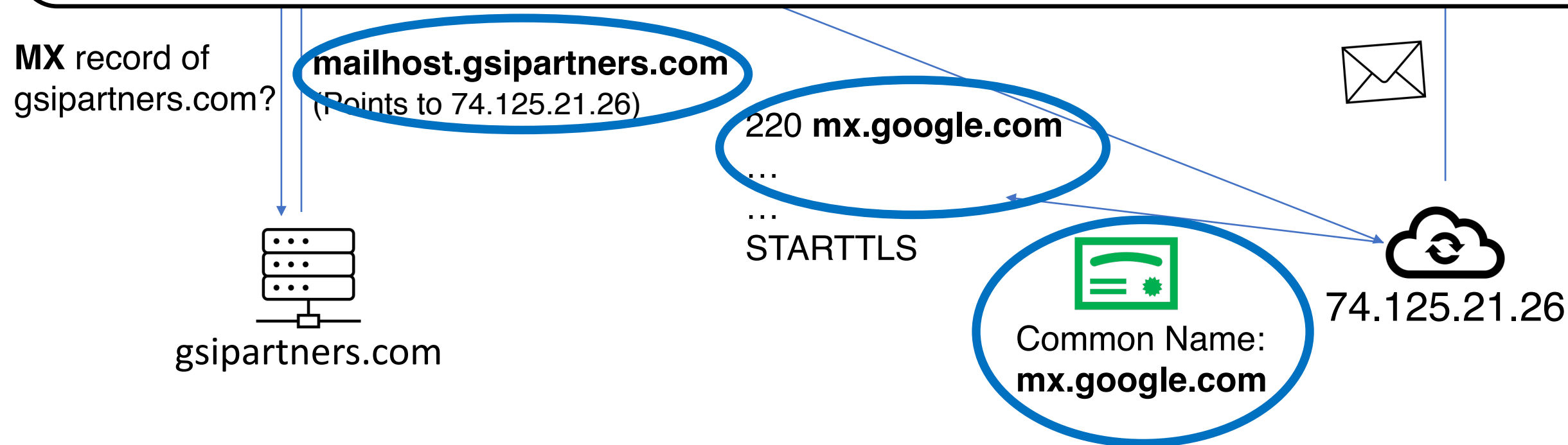
Determine the Mail Service Provider

Video Here



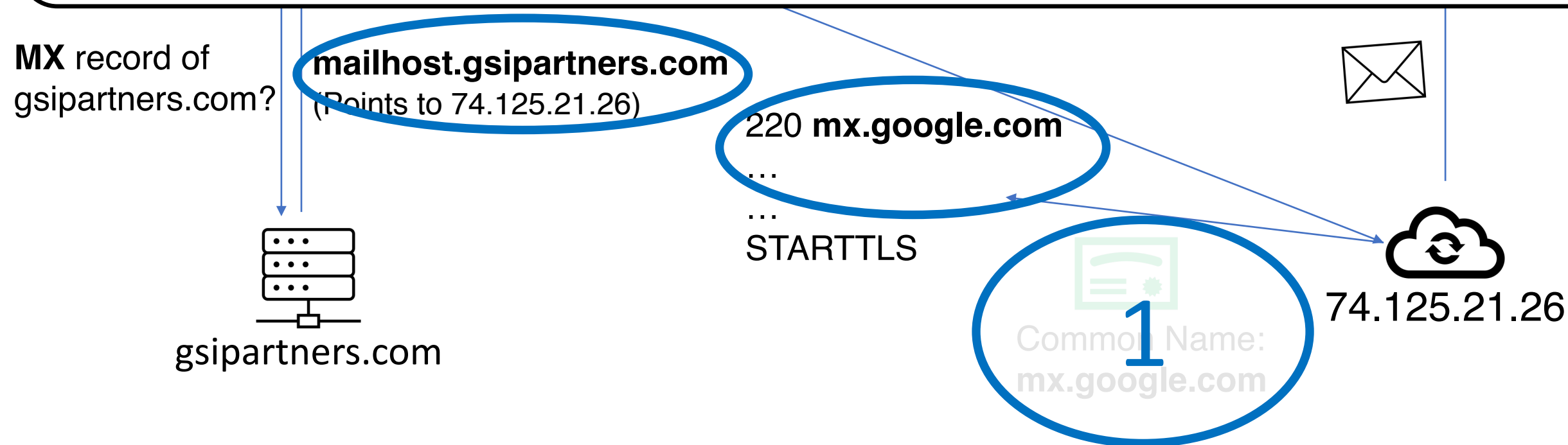
Determine the Mail Service Provider

Key Insight: prioritize reliable data source



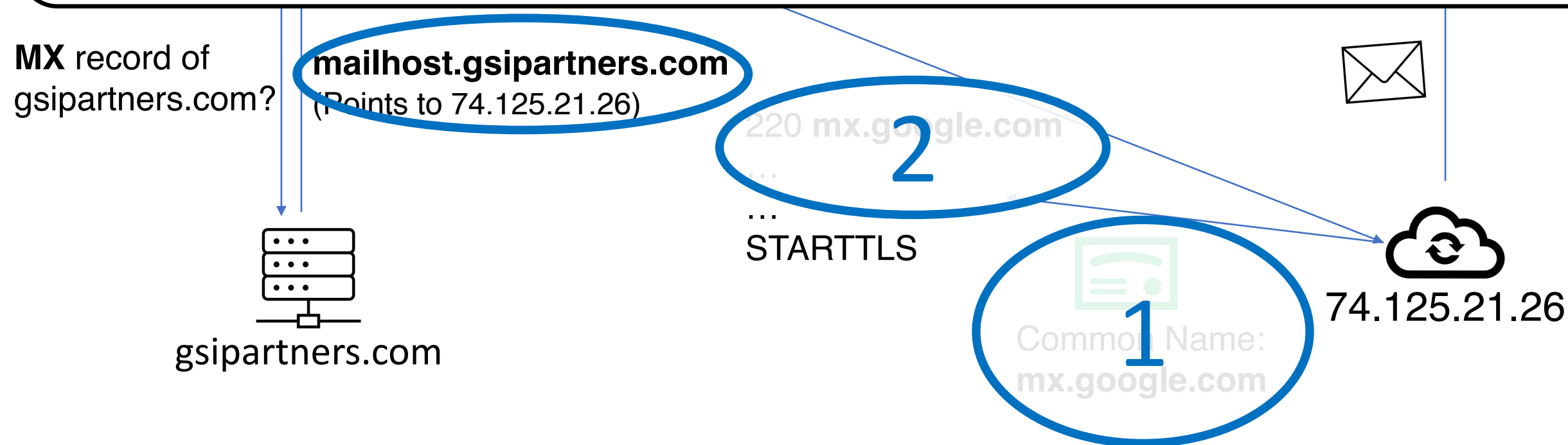
Determine the Mail Service Provider

Key Insight: prioritize reliable data source



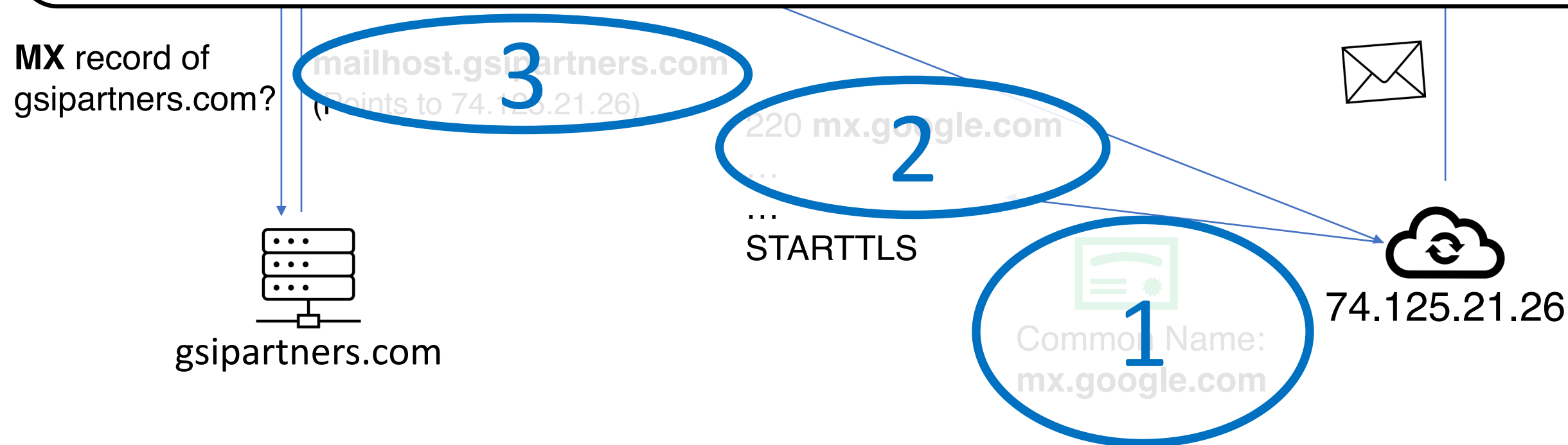
Determine the Mail Service Provider

Key Insight: prioritize reliable data source



Determine the Mail Service Provider

Key Insight: prioritize reliable data source



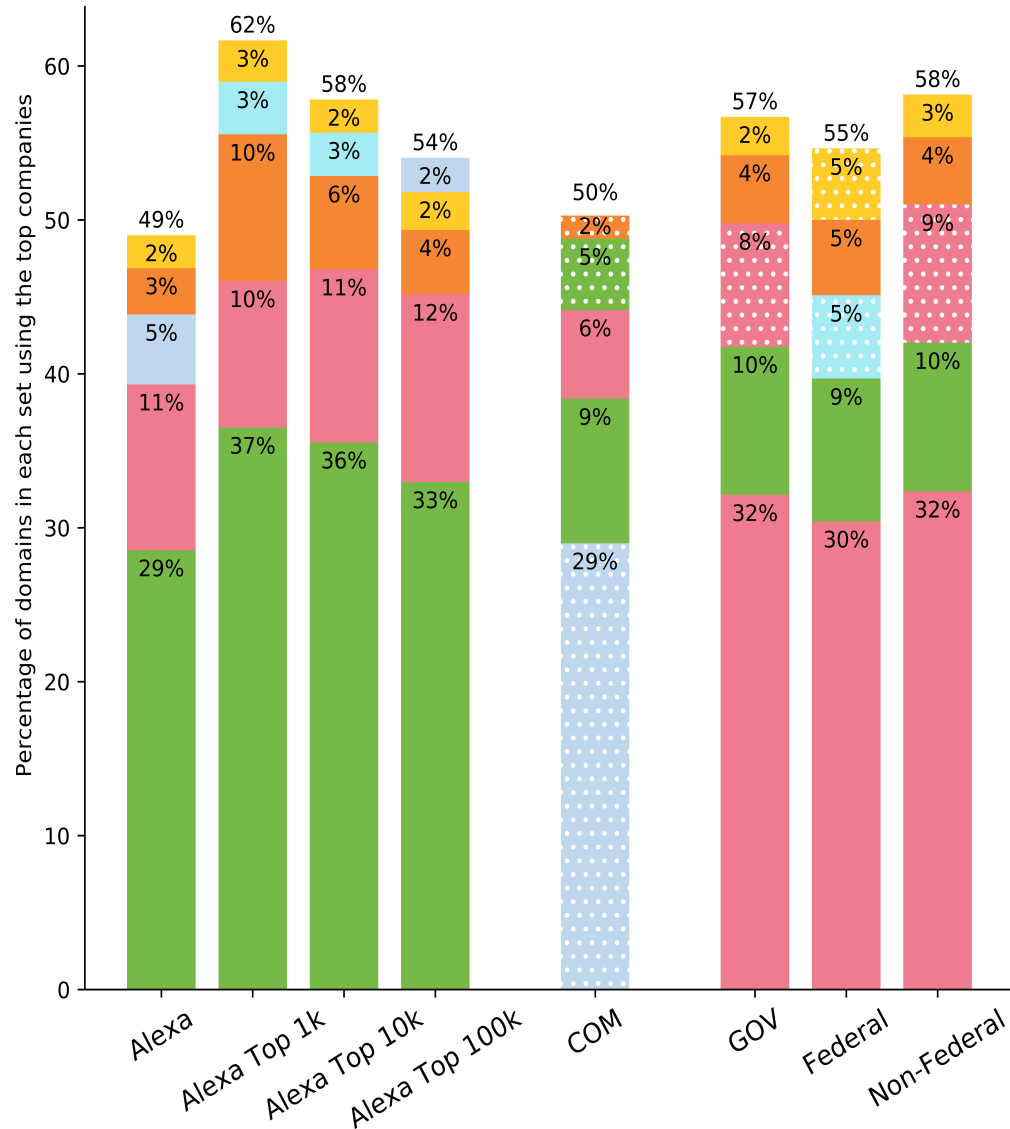
Contributions

- Methodology for mapping domains to mail service providers
- **Characterization** of mail service provider usage
 - Current landscape
 - Longitudinal trends
 - Provider preference by country

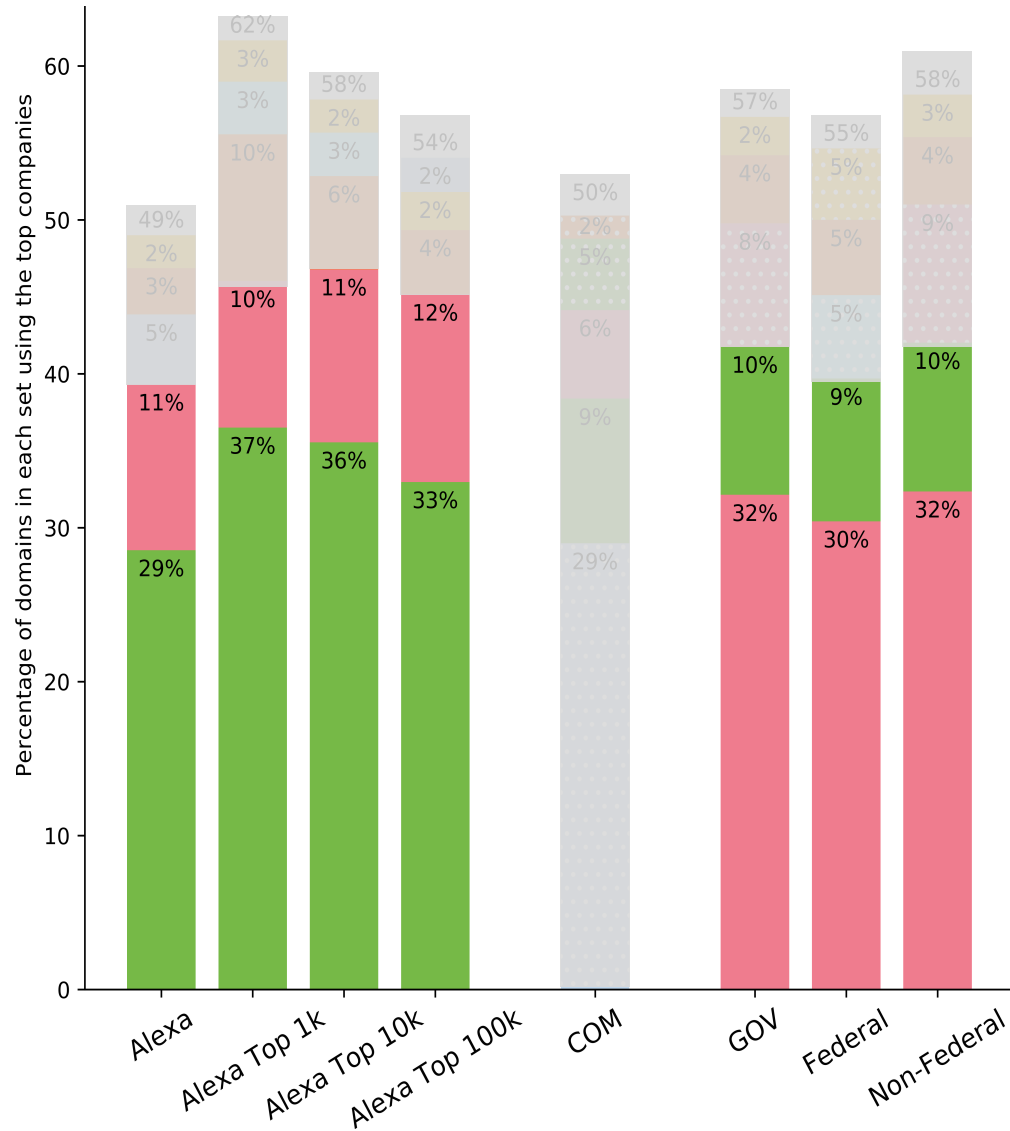
Datasets

- Timespan: June 2017 - June 2021 w/ nine snapshots
- Three sets of target domains w/ MX records
 - Alexa top 1m domains → 93k **stable** Alexa domains Popular
 - 1m randomly sampled .com domains → 580k **stable** .com domains Random
 - All .gov domains → 3.5k **stable** .gov domains Restricted TLD
- Two external dataset sources:
 - OpenINTEL: MX / MX resolution data
 - Censys: IP scan data

Current Landscape

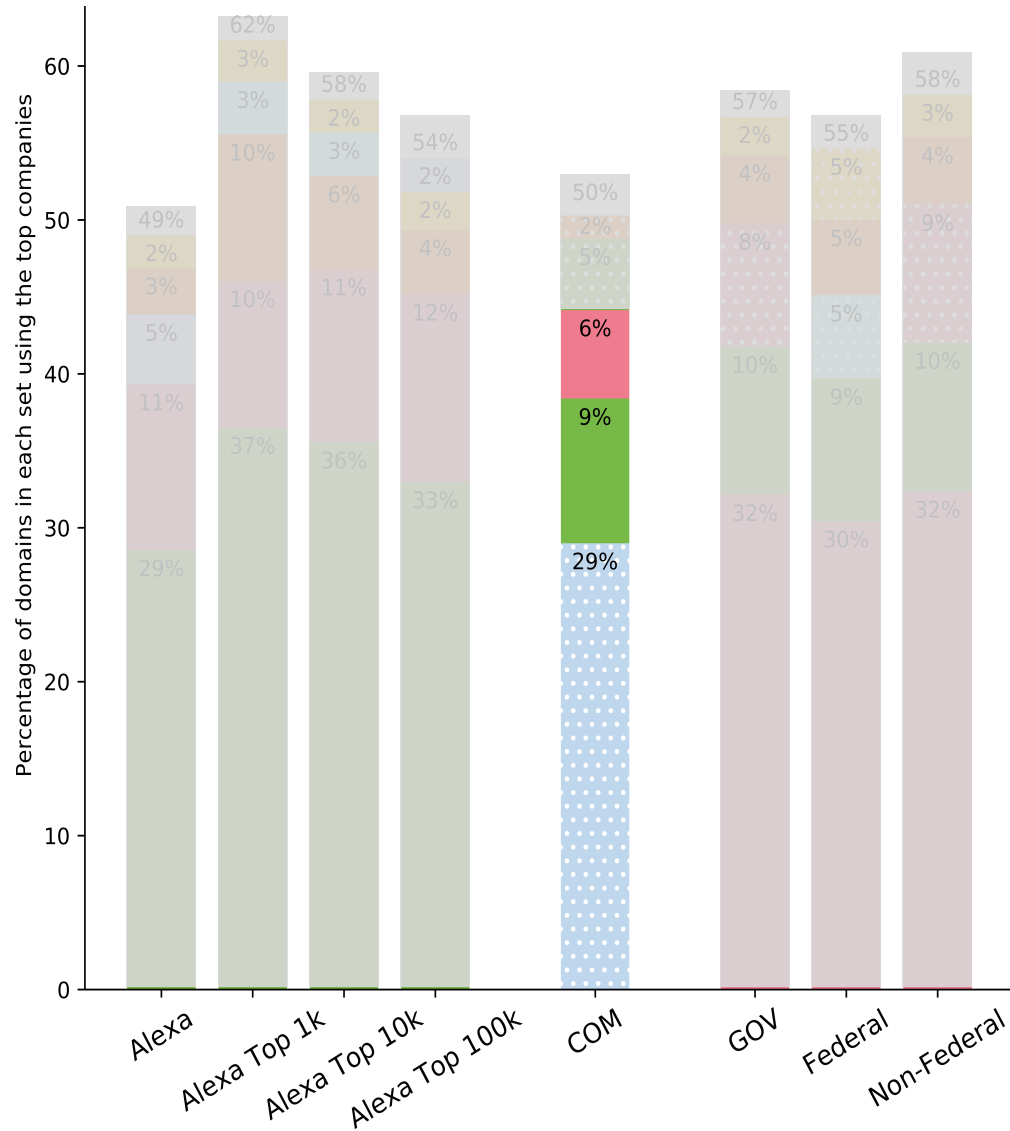


Current Landscape



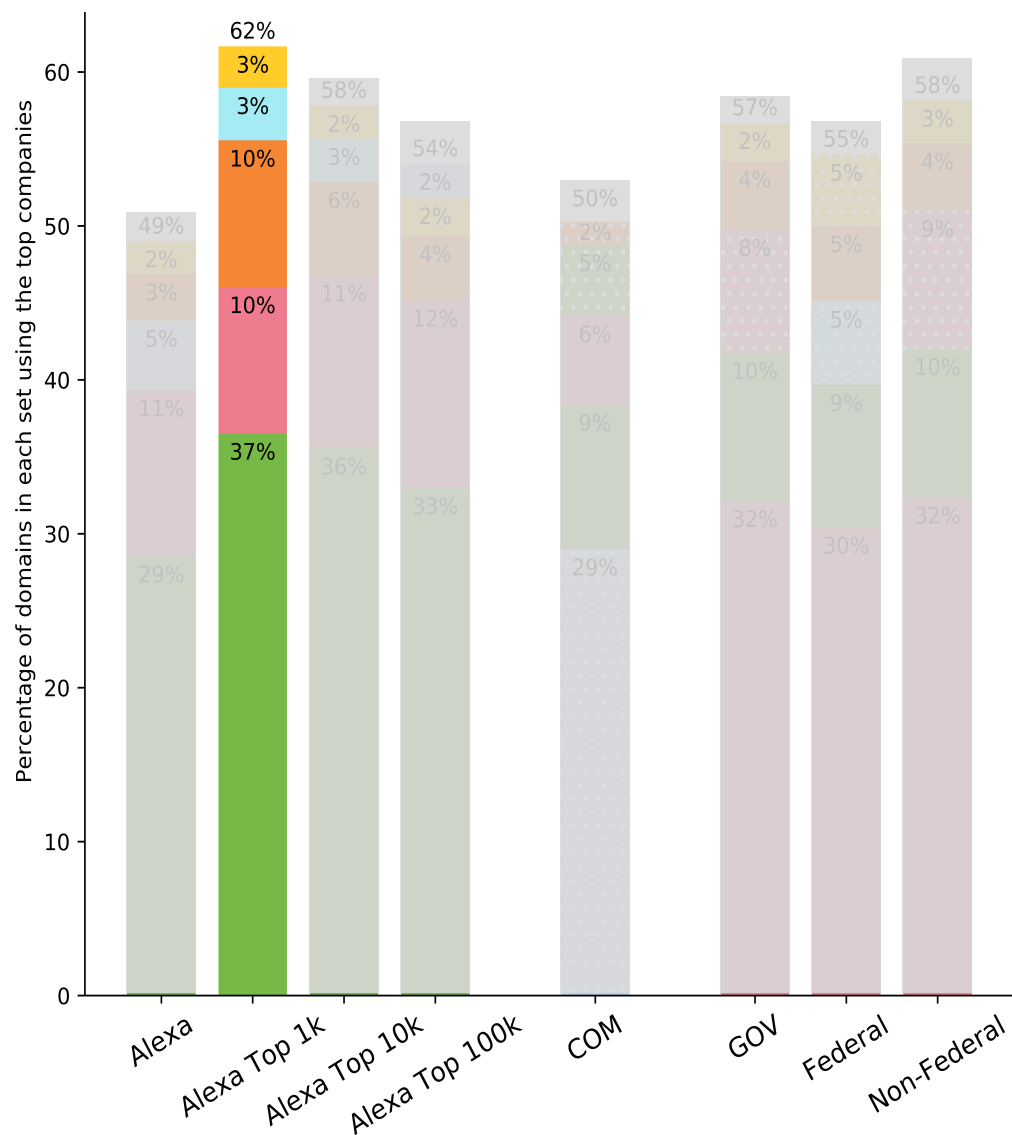
- Google and Microsoft are dominant for Alexa and .gov

Current Landscape



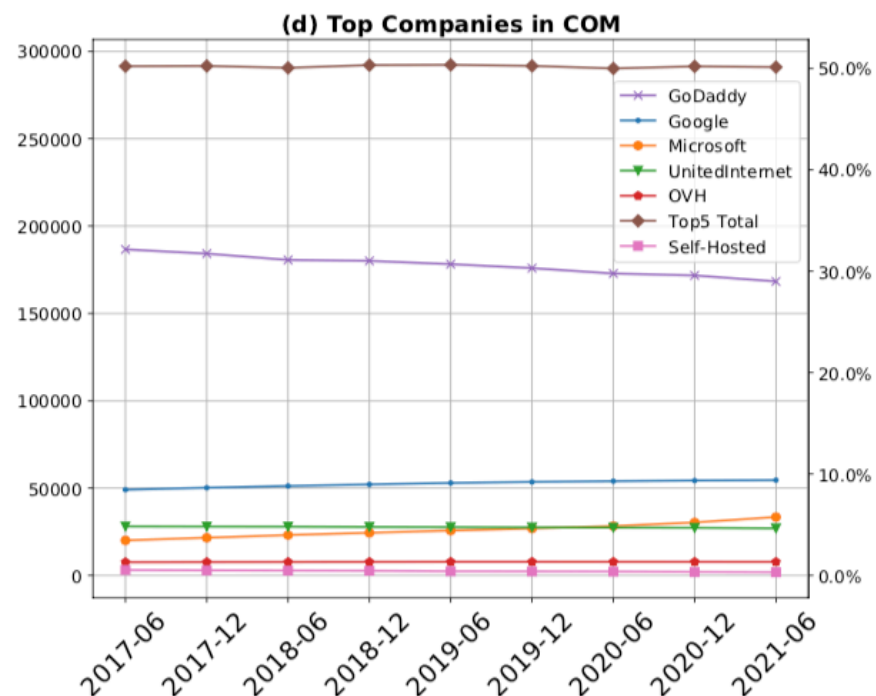
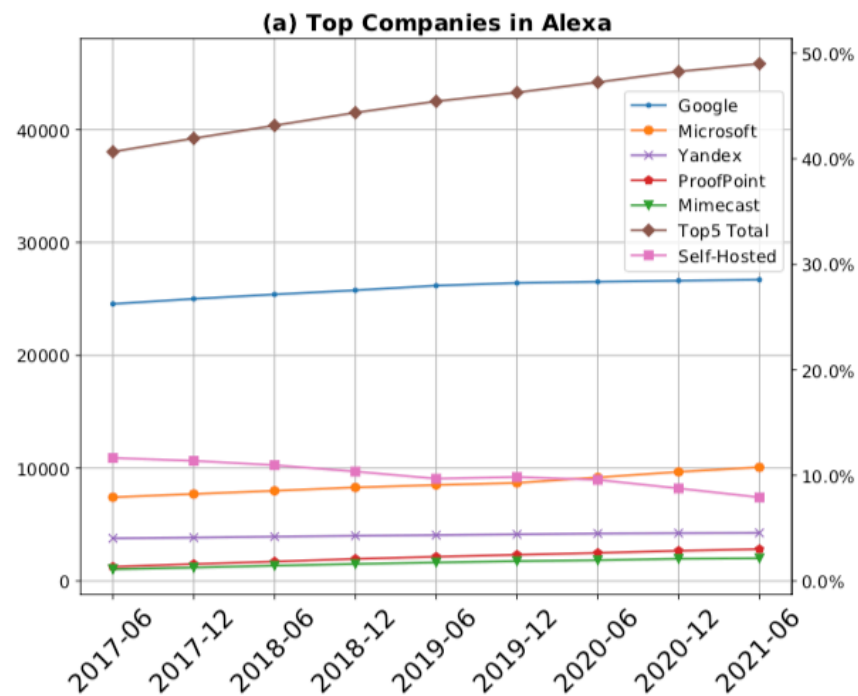
- Google and Microsoft are dominant for Alexa and .gov
- GoDaddy has the domain market share for .com

Current Landscape

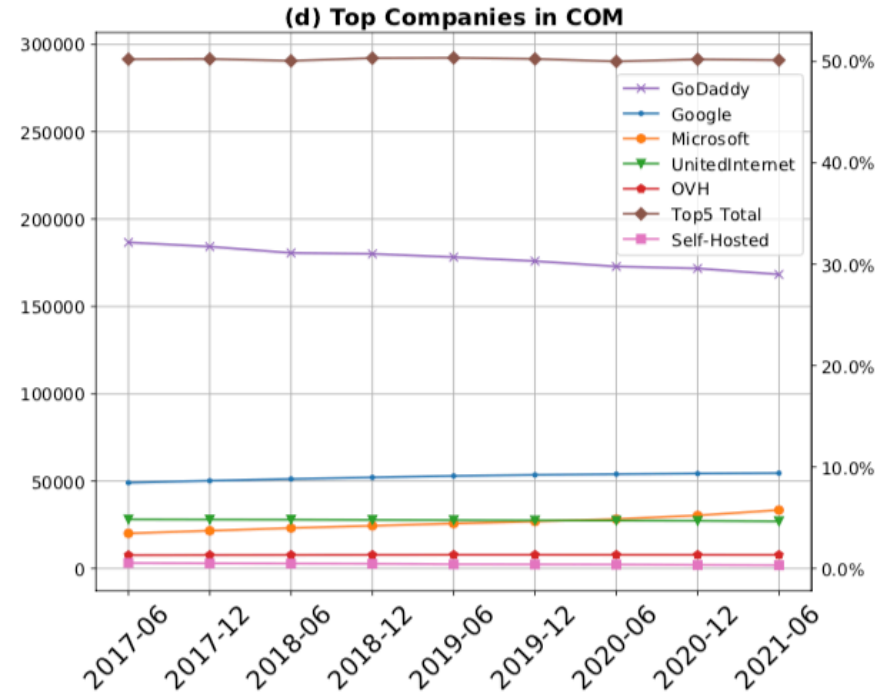
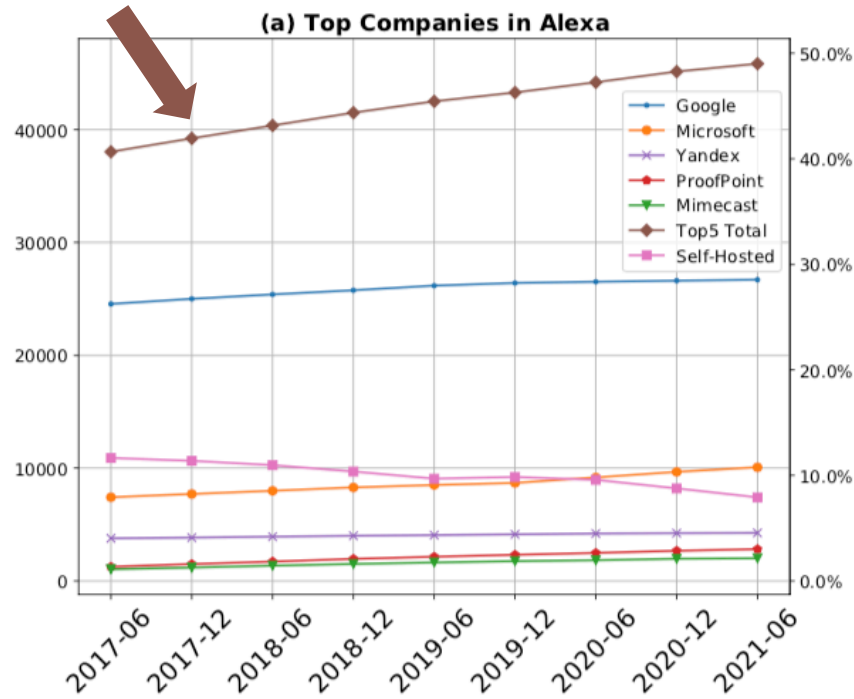


- Google and Microsoft are dominant for Alexa and .gov
- GoDaddy has the domain market share for .com
- Higher-level of centralization among Alexa top 1k

Longitudinal Trends

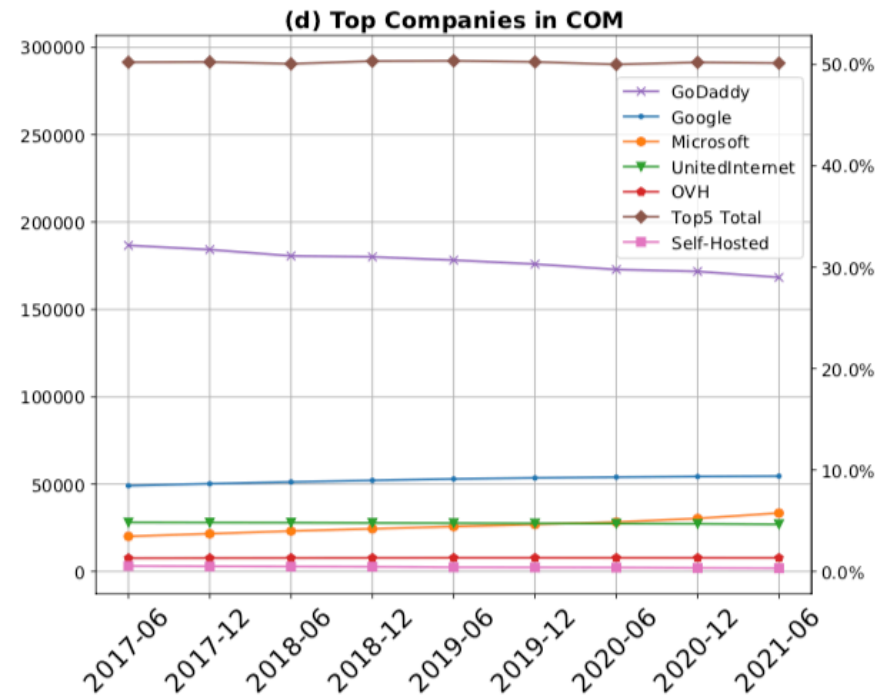
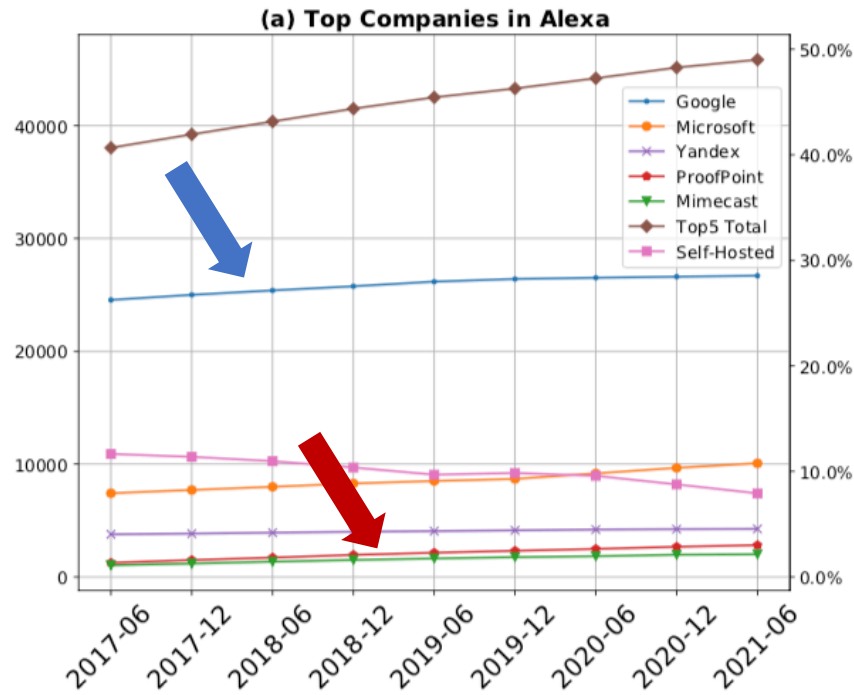


Longitudinal Trends



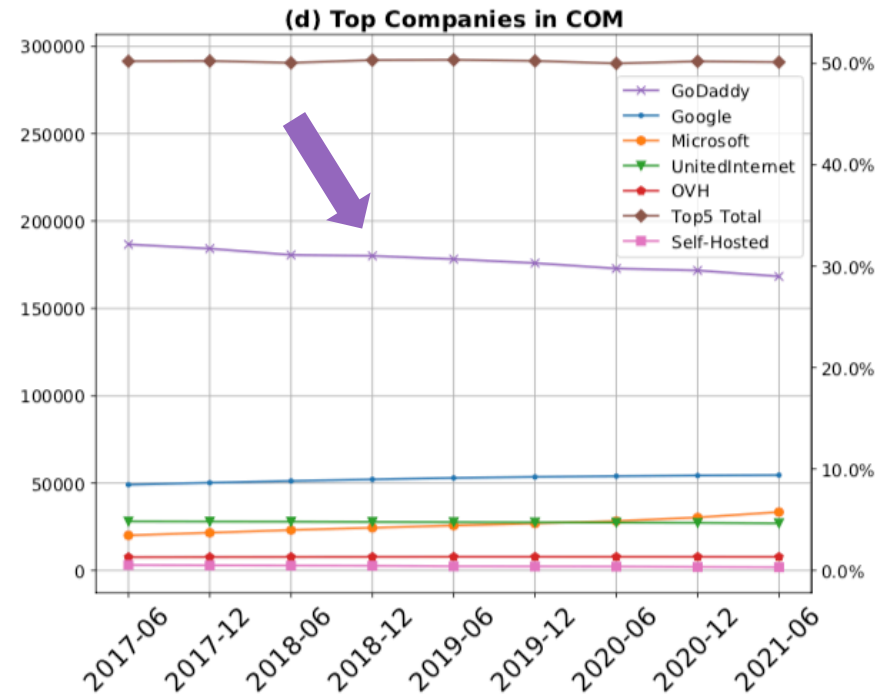
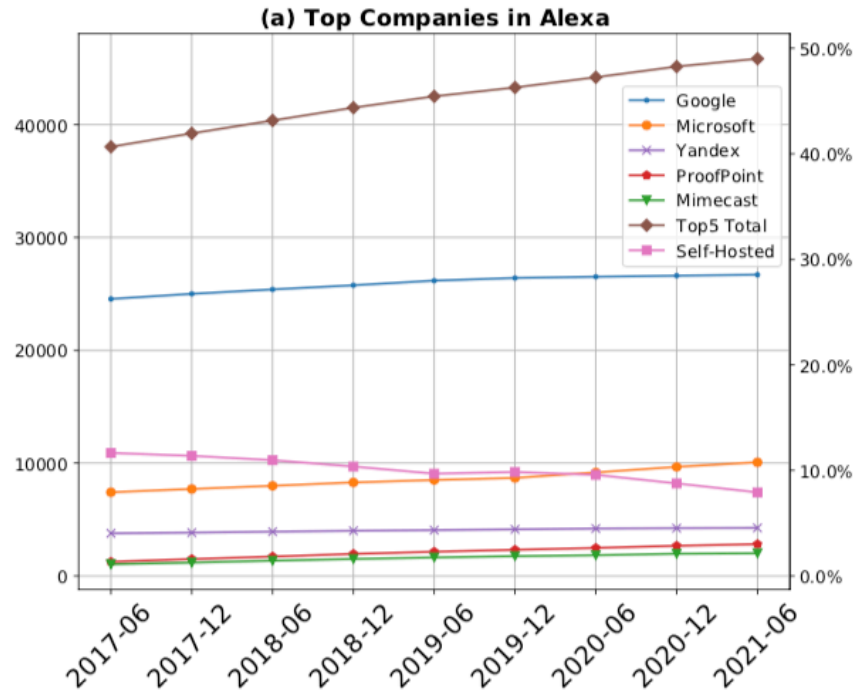
- In general, we see **increasing centralization** among the top 5 providers over 4 years

Longitudinal Trends



- In general, we see **increasing centralization** among the top 5 providers over 4 years
 - Increased popularity of **hosting companies** (e.g., Google) and **security companies** (e.g., Proofpoint)

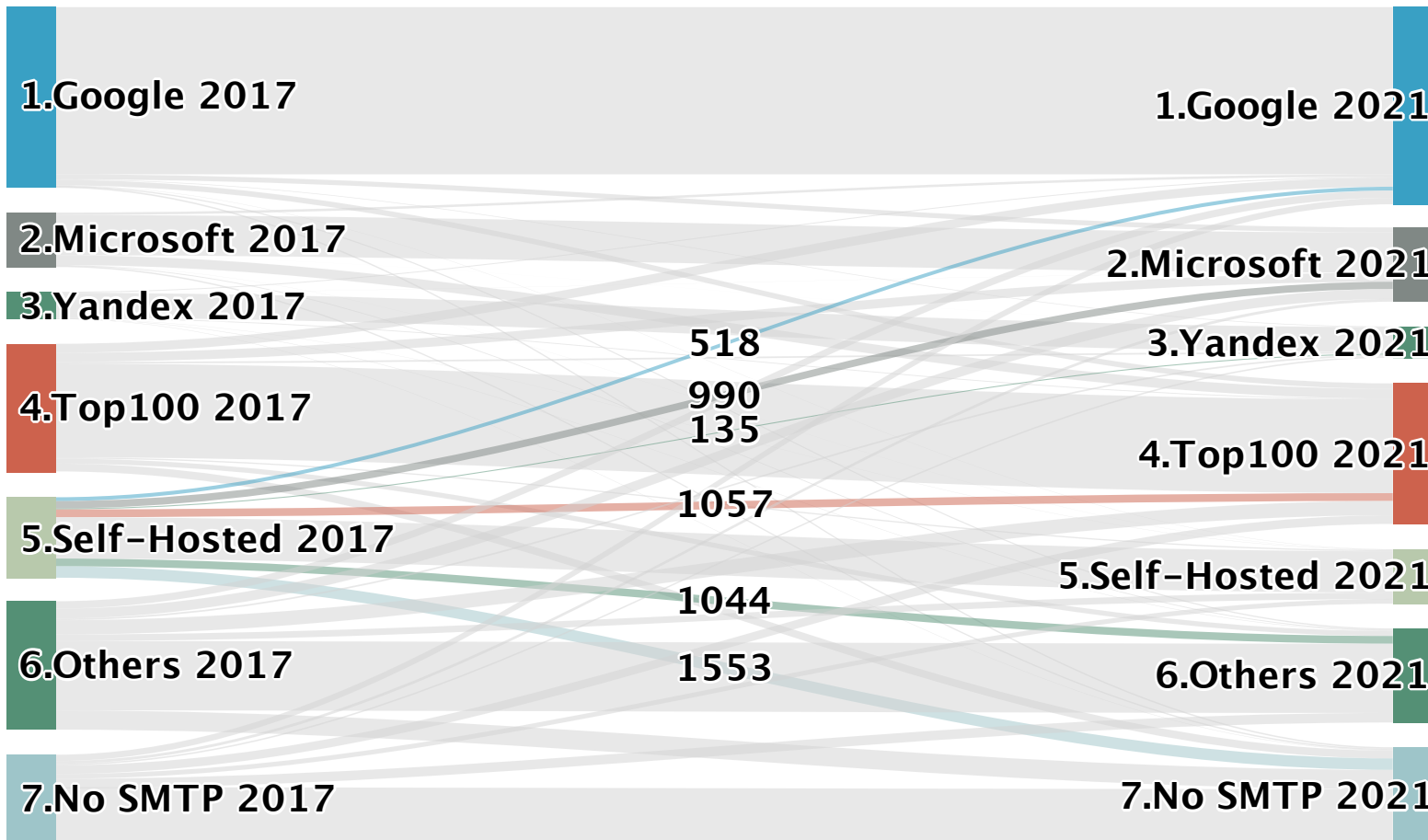
Longitudinal Trends



- In general, we see **increasing centralization** among the top 5 providers over 4 years
 - Increased popularity of **hosting companies** (e.g., Google) and **security companies** (e.g., Proofpoint)
- **Fewer domains** use email services provided by web hosting companies like GoDaddy

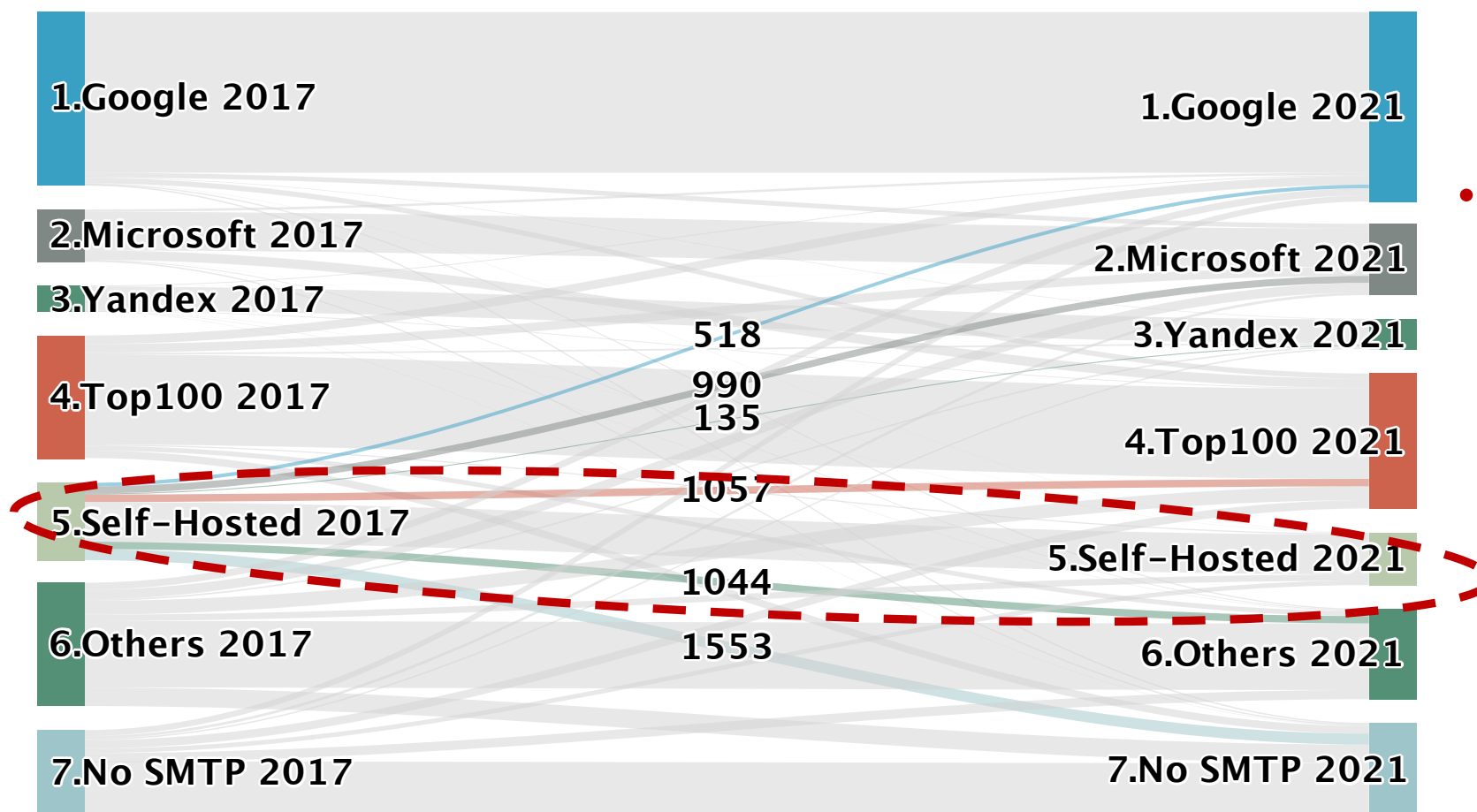
Self-Hosted Domains

Churn in Self-Hosted Domains (2017 to 2021)



Self-Hosted Domains

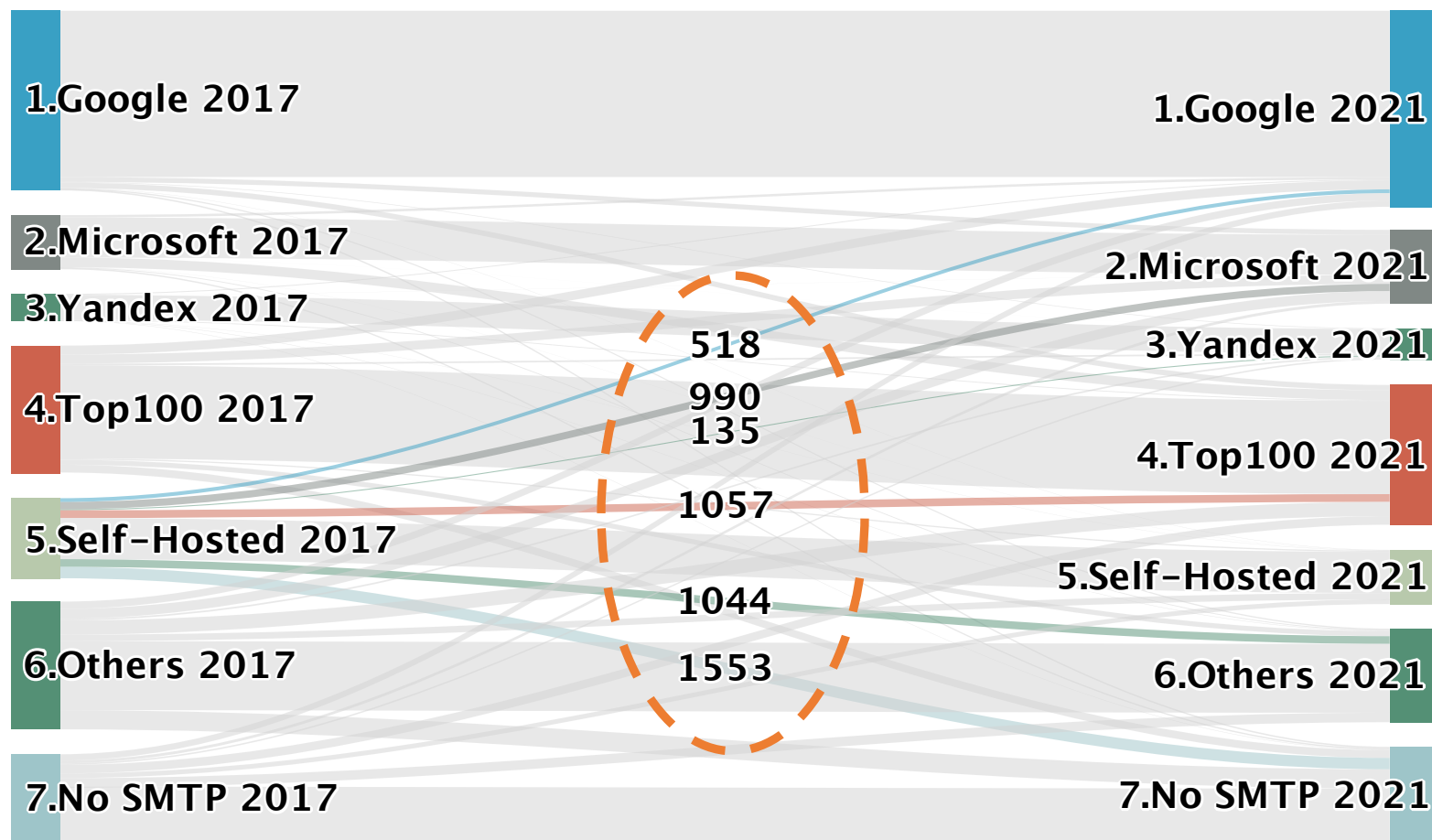
Churn in Self-Hosted Domains (2017 to 2021)



- Fewer domains are hosting their own mail services

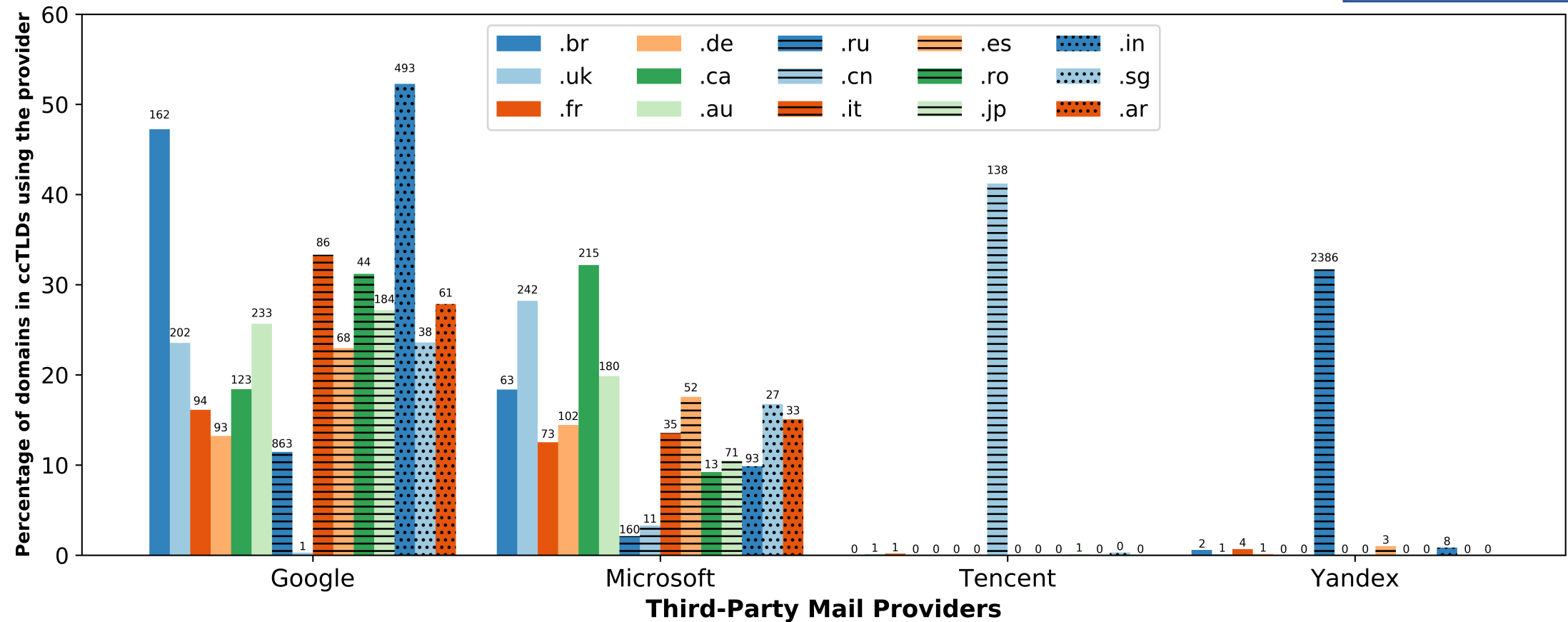
Self-Hosted Domains

Churn in Self-Hosted Domains (2017 to 2021)

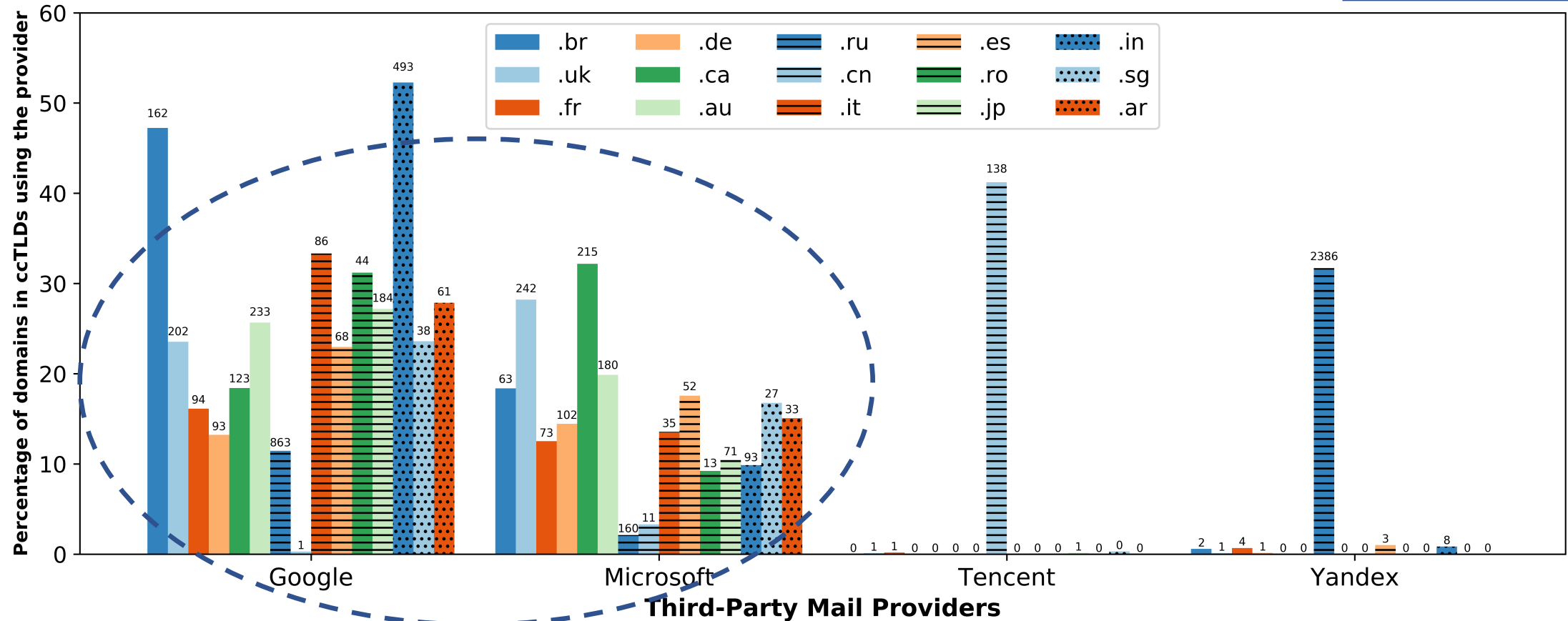


- Fewer domains are hosting their own mail services
- Among self-hosted domains that switched providers, more than 25% of them switched to Google/Microsoft

Mail Provider by Country

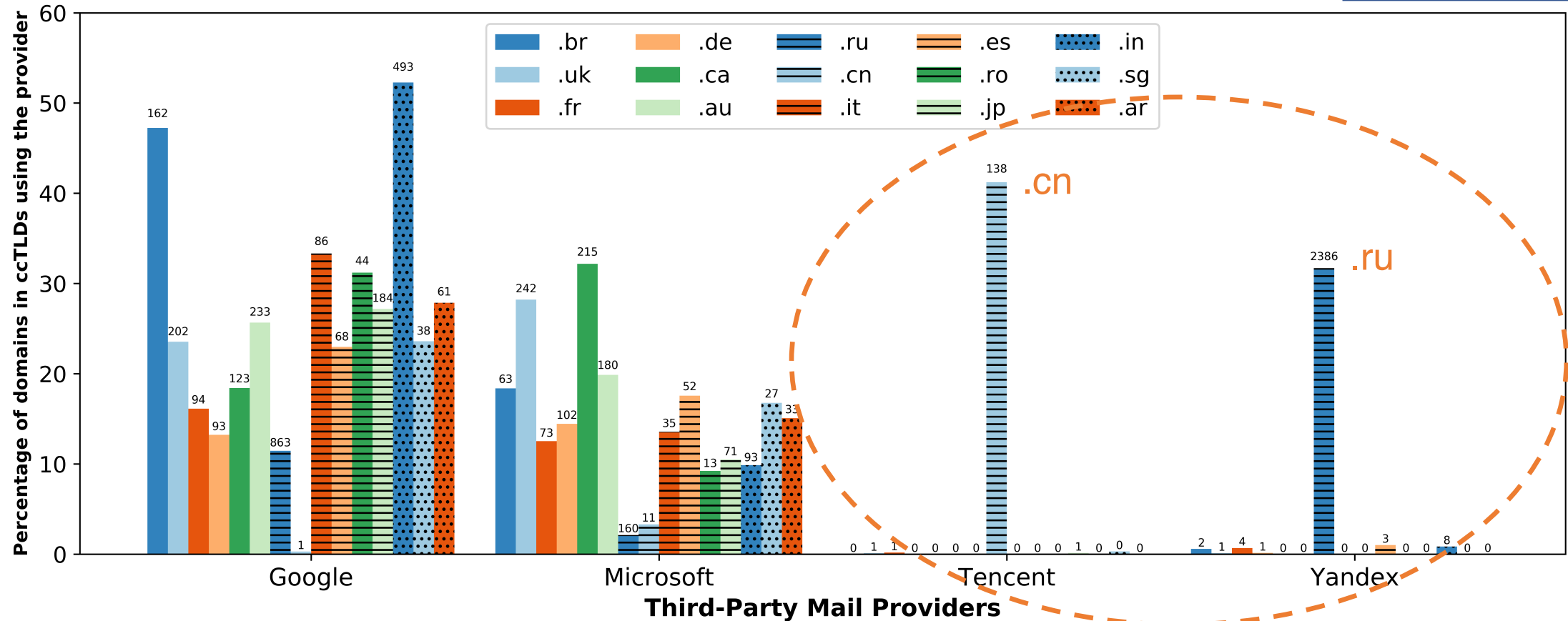


Mail Provider by Country



1. Google/Microsoft are widely used across the globe (including Russia but not China)

Mail Provider by Country



1. Google/Microsoft are widely used across the globe (including Russia but not China)
2. Tencent/Yandex primarily serve domestic markets

Summary

- **Methodology** for mapping domains to mail service providers
 - Details in the paper
 - **Code:** https://github.com/ucsdsysnet/mx_inference
- **Characterization** of mail service provider usage
 - Current landscape
 - Longitudinal trends
 - Provider preference by country
- Got questions? Email **Alex Liu** at e7liu@eng.ucsd.edu