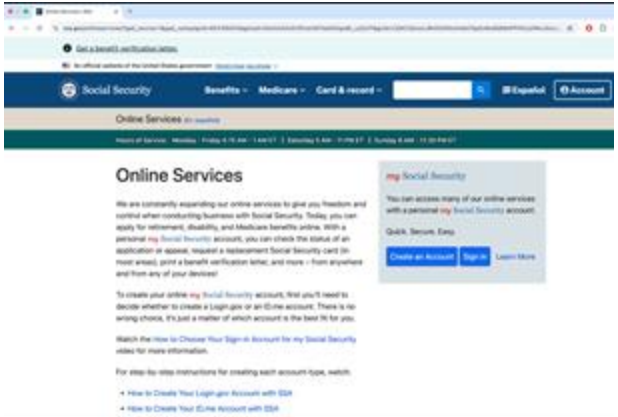
When Consolidation is Not a Choice

Rashna Kumar
Research Fellow, Internet Society . PhD Candidate, Northwestern University

Mentors: Marinho Barcellos, Amreesh Phokeer, Fabián E. Bustamante

Websites Depend on Multiple Services



- Domain Name System (DNS)
  - Translate domain names to IP addresses

CLOUDFLARE

amazon web services Route 53

- Hosting
  - Serve content

    Often from CDN servers for faster, reliable access

fastly Akamai

- HTTPS? Certificate Authority (CA)
  - Verify validity of SSL certificates to secure websites

digicert | Symantec

These Services are Consolidating…

A few companies now operate much of the web's foundational infrastructure

Consolidation is a Problem for Resilience

Amazon reveals cause of AWS outage that took everything from banks to smart beds offline

AWS explains in a lengthy post how a bug in automation software brought down thousands of sites and applications

Signal, Snapchat, Roblox, Duolingo and Ring doorbells were some of the 2,000 companies affected by this week's AWS outage, according to Downdetector. Photograph: Anushree Fadnavis/Reuters

Cloudflare's CTO apologizes after error takes huge chunk of the internet offline — 'we failed our customers and the broader internet'

News By Luke James published November 18, 2025

CTO blames bot mitigation bug triggered by routine config change.

(Image credit: Getty / Smith Collection/Gado)

Cloudflare has confirmed that a bug in one of its core services caused a major outage on Tuesday, taking large portions of the internet offline and affecting traffic to services including X, ChatGPT, and, ironically, Downdetector. The company's CTO, Dane Knecht, posted a public apology shortly after services were restored, calling the incident "unacceptable" and attributing the disruption to a routine configuration change that triggered a crash in its bot mitigation layer.

Microsoft Azure Outage (Oct 29 2025): Root Cause, Impact and Technical Analysis

Ismail Kovvuru Follow 9 min read · Oct 31, 2025

A deep technical breakdown of the Microsoft Azure outage on Oct 29 2025. Learn how an Azure Front Door configuration change disrupted global services including Microsoft 365, Xbox Live and airline systems plus insights on root cause, mitigation and lessons for teams.

Prior Work on Consolidating Internet

Analyzing Third Party Service Dependencies in Modern Web Services: Have We Learned from the Mirai-Dyn Incident?

Aqsa Kashaf, Carnegie Mellon University, akashaf@andrew.cmu.edu
Vyas Sekar, Carnegie Mellon University, vsekar@andrew.cmu.edu
Yuvraj Agarwal, Carnegie Mellon University, yuvraja@cs.cmu.edu

Abstract

Many websites rely on third parties for services (e.g., DNS, CDN, etc.). However, it also exposes them to shared risks from attacks (e.g., Mirai DDoS attack [26]) or cascading failures (e.g., GlobalSign revocation error [21]). Motivated by such incidents, we analyze the prevalence and impact of third-party dependencies, focusing on three critical infrastructure services: DNS, CDN, and certificate revocation checking by CA. We analyze both direct (e.g., Twitter uses Dyn) and indirect (e.g., Netflix uses Symantec as CA which uses Verisign for DNS) dependencies. We also take two snapshots (in 2016 and 2020 to understand how the dependencies evolved. Our key findings are (1) 89% of the Alexa top-100K websites critically depend on third-party DNS, CDN, or CA providers; i.e., if those providers go down, these websites could suffer service disruption; (2) the use of third-party services is concentrated, and the top 3 providers of CDN, DNS, or CA services can affect 50%-70% of the top-100K websites; (3) indirect dependencies amplify the impact of popular CDN and DNS providers by up to 25X, and (4) some third-party dependencies and concentration increased marginally between 2016 to 2020. Based on our findings, we derive key implications for different stakeholders in the web ecosystem.

Each at its Own Pace: Third-Party Dependency and Centralization Around the World

RASHNA KUMAR, Northwestern University, USA
SANA ASIF, Northwestern University, USA
ELISE LEE, Northwestern University, USA
FABIÁN E. BUSTAMANTE, Northwestern University, USA

We describe the results of a large-scale study of third-party dependencies around the world based on regional top-500 popular websites accessed from vantage points in 50 countries, together covering all inhabited continents. This broad perspective shows that dependencies on a third-party DNS, CDN or CA provider vary widely around the world, ranging from 19% to as much as 76% of websites, across all countries. The critical dependencies of websites – where the site depends on a single third-party provider – are equally spread ranging from 5% to 60% (CDN in Costa Rica and DNS in China, respectively). Interestingly, despite this high variability, our results suggest a highly concentrated market of third-party providers: three providers across all countries serve an average of 92% and Google, by itself, serves an average of 70% of the surveyed websites. Even more concerning, these differences persist a year later with increasing dependencies, particularly for DNS and CDNs. We briefly explore various factors that may help explain the differences and similarities in degrees of third-party dependency across countries, including economic conditions, Internet development, economic trading partners, categories, home countries, and traffic skewness of the country's top-500 sites.

Formalizing Dependence of Web Infrastructure

Rumaisa Habib, Stanford University, Stanford, CA, USA
Kimberly Ruth, Stanford University, Stanford, CA, USA
Gautam Akiwate, Stanford University, Stanford, CA, USA
Zakir Durumeric, Stanford University, Stanford, CA, USA

ABSTRACT

Hosting Industry Centralization and Consolidation

Luciano Zembruzki, Raffaele Sommese, Lisandro Zambenedetti Granville, Arthur Selle Jacobs, Mattijs Jonker, Giovane C. M. Moura

Of Choices and Control – A Comparative Analysis of Government Hosting

Clouding up the Internet: how centralized is DNS traffic becoming?

Giovane C. M. Moura, Sebastian Castro, Wes Hardaker, Maarten Wullink, Cristian Hesselman

ABSTRACT

The research community has been tracking this growing consolidation for years yet it keeps accelerating

Governments are Increasingly Relying on E-Services

Tax filing

Healthcare

Immigration

Licensing

Voting

Importance of digital government:

- Federal websites in the US draw nearly two billion visits every month
- In the Asia-Pacific region, 77 percent of citizens primarily access government services through digital platforms

Prior Work Shows Governments are Shaped by Same Consolidation Trends

Hosting markets are already consolidated for government services:

- Cloudflare hosts e-services for 49 governments
- And serves up to 72% of all government web bytes for one country

But Do Governments Have a Choice? Are They Opting to Be Consolidated?

Constraint-Driven (Structural)

Potential reasons:

Few or no domestic infrastructure options

Consequence:

Shared dependency across government and commercial web

Choice-Driven (Strategic)

Potential reasons:

Perceived reliability, cost, procurement rules, vendor ties

Consequence:

Government chooses consolidation, where commercial web shows options

How We Determine Whether Consolidation Is a Choice

To distinguish what the environment forces from what governments choose

Block by country (structural baseline)

- Each country forms a block with shared infrastructure, regulation, market conditions
- This fixes the structural environment.

Within-block groups

- Treated: Government domains
- Reference: Commercial domains from the same country, stratified by popularity:
  - Top-1k: large/global actors
  - Mid-1k: mid-tier actors
  - Bottom-1k: small/local actors

Inference: Compare consolidation vs. decentralization across groups within a block

- Similar patterns → structural forces
- Divergent patterns → strategic choices by governments

Measurement Design: Collecting Data

Scope and Inputs

Countries: 61, spanning every continent (over 82% of the world's Internet population)

Domains Per Country:

- All Government sites
- 3K Commercial sites (CrUX)

Building Dataset

Government domains: Collected from official sources

Commercial Domains:

- Split CrUX domains per country into Top / Middle / Bottom stratas
- From each strata, collect 1K sites using systematic step sampling (uniform +popularity-aware coverage)

Service Extraction

Connect to country-local VPN and collect for each landing page:

- All resources and their hosting orgs
- Authoritative DNS nameservers
- Certificate Authorities from TLS handshake

Measurement Design: Analyzing Data

Provider Concentration (HHI)

For each strata:

- Resample domains with replacement (same sample size)
- Repeat 5,000 bootstrap iterations

For each country × service × strata:

- Obtain distribution of HHI values

Compare Consolidation within blocks

For each bootstrap iteration, compute ΔHHI

$\Delta HHI = HHI_{commercial} - HHI_{government}$

Produce bootstrap distribution of ΔHHI per (country × service × strata)

Check statistical significance of ΔHHI (p-values)

Compare Patterns across blocks

Structural Consolidation: ΔHHI not significant (and Government provider diversity resembles commercial sites)

Strategic Consolidation: ΔHHI significant (Clear separation between government and commercial sites)

Government Consolidation Across Services

These maps show how consolidated government web infrastructure is across services.

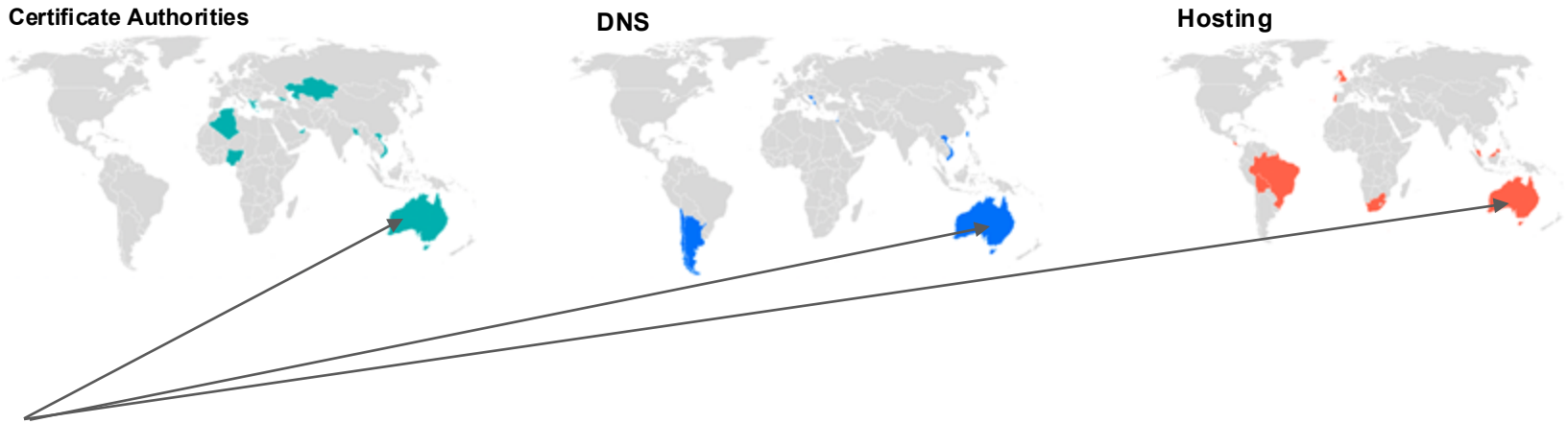
Colors reflect HHI thresholds commonly used in antitrust

Governments worldwide are highly consolidated on Certificate Authorities

Hosting is overall more moderate, with a fewer cases of high consolidation

Structural Government Consolidation Across Services

These countries show similar moderate-to-high consolidation across government and commercial strata

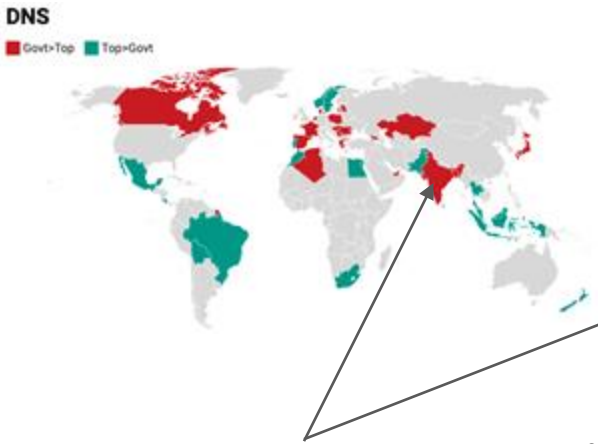


Australia's consolidation patterns reflect structural forces rather than strategic choices

- Government and commercial sites rely on a similar set of certified providers across all services
- Australian government follows a Cloud First strategy and a multi-vendor model, avoiding reliance on a single national provider

Strategic Government Consolidation Across Services

These maps show where government consolidation departs from commercial patterns



Indian government sites show a huge reliance on a domestic CA eMudhra, reducing reliance on global CAs

DNS and Hosting: >85% of Indian government sites rely on the National Informatics Centre (NIC)

Strategic Government Consolidation Across Services

These maps show where government consolidation departs from commercial patterns

Kazakhstan:

DNS: Dominated by Hoster.kz and state-linked providers

Hosting: Strong reliance on Kazakhtelecom and domestic infrastructure

Strategic Government Consolidation Across Services

These maps show where government consolidation departs from commercial patterns

Algeria:

DNS: All Government sites rely on Cloudflare DNS

Hosting: Government sites highly consolidated under Telecom Algeria (>80%)

Our Framework Reveals

When constraints drive outcomes and when choices create threats to resilience

Our framework separates what governments *must* do from what they *choose* to do

Building a More Resilient Internet

Support where consolidation is structural

Invest to build local DNS, hosting and CA capacity

Intervene where consolidation is strategic

Policy nudges can reduce unnecessary dependency

Internet Society

Can help strengthen Internet Resilience

By local infrastructure development and policy engagement

Ongoing Work

- Clustering countries by regulatory and market characteristics
- Assessing implications of strategic consolidation for resilience and sovereignty
- Analyzing on-path dependencies

More insights to come as we deepen the analysis!

Thank you!

Happy to *get in touch for any questions:*

rashnakumar2024@u.northwestern.edu

https://sites.northwestern.edu/rashnakumar/