

The Handoff Behind Your Click

Hidden Maps · Episode 4 — Who Carries the Internet

Troy Altus

2026-06-01

Table of contents

1	A note before the numbers	1
2	Abstract	2
3	Introduction: You Never See This Map	2
4	Data: The CAIDA AS Graph	3
5	Methods (without the math wall)	3
6	Results	4
6.1	Unequal connections (the hub story)	4
6.2	Small-world structure	4
6.3	Who sits on the chokepoints?	4
6.4	Communities: neighborhoods of the internet	5
6.5	Robustness: random outages vs. hitting the backbones	5
7	Discussion: What this does and does not claim	5
8	Conclusion	6

1 A note before the numbers

New to this series? Read [Hidden Maps](#) first. It explains the whole toolkit in plain language — dots, lines, hubs, chokepoints, and the difference between random bad luck and hitting the right failure point.

This episode is the same story applied to the internet — not websites and apps, but the independent networks underneath them.

2 Abstract

When a webpage loads, your data is handed from network to network — your home ISP, a regional carrier, a global backbone, a cloud provider — each step governed by business contracts you never see. This study maps that hidden layer using the June 2026 CAIDA AS-relationships snapshot (CAIDA 2025): 78,144 Autonomous Systems and 571,399 routing links in the largest connected component. The internet’s control plane is a classic heavy-tailed network — a few ASes hold enormous connection counts, and betweenness concentrates on transit providers that bridge communities. Under random AS withdrawal the map stays connected surprisingly long; under targeted removal of high-betweenness nodes, the largest component collapses after roughly 0.8% of nodes are gone. The same robust-yet-fragile pattern found in flights, power lines, and city streets appears here at planetary scale.

3 Introduction: You Never See This Map

You notice the internet when it fails: a Cloudflare outage turns half the web gray, a submarine cable cut slows a continent, a peering dispute between two carriers adds latency you cannot diagnose. What you rarely see is the map those failures trace.

That map is not geographic. It is contractual. The internet is stitched together from roughly 78,144 **Autonomous Systems** — independent routing domains run by ISPs, universities, governments, and hyperscale cloud companies. Each AS announces: *I know how to reach these addresses. Send me traffic, and I will forward it according to my contracts with my neighbors.*

An Autonomous System is not a company exactly, and not a physical cable. It is a unit of routing policy — a numbered participant in the global Border Gateway Protocol (BGP) conversation. When you click a link, packets hop across AS boundaries the way flights connect through hub airports. The difference is that nobody designed this graph from above. It grew from peering handshakes, customer contracts, and decades of commercial negotiation (Luckie et al. 2013).

This is **Episode 4** of the Hidden Maps series:

Episode	System	What breaks in real life
Primer	The toolkit	—
1 · Flights	US airports	Delays that ripple nationwide
2 · The Grid	Transmission lines	Blackouts from one failed path
3 · The Streets	Pittsburgh roads	One closed bridge, everyone’s late
4 · The Internet	AS routing	Outages, cables, peering fights

Three questions, same as every episode:

1. **Shape:** Is connection count wildly unequal — a few giants, a long tail of small networks?
2. **Chokepoints:** Which ASes sit on the most through-paths, regardless of brand fame?
3. **Breakage:** Random outages vs. deliberate removal of the most central ASes — how fast does the map split?

4 Data: The CAIDA AS Graph

The analysis uses CAIDA’s **AS-relationships serial-2** dataset for 1 June 2025 (CAIDA 2025) — a monthly snapshot inferred from public BGP routing tables collected worldwide. Each line records a business relationship between two AS numbers:

- **Peer–peer (0):** two networks of comparable size exchange traffic as equals
- **Customer–provider (-1):** one network pays another to carry its traffic

For structural analysis, we treat the graph as **undirected**: if two ASes route traffic between them, they are connected. The raw file contains 571,399 relationship records; after building a simple graph and keeping the **largest connected component**, we analyze 78,144 ASes linked by 571,399 edges.

Table 1: Internet AS graph statistics (June 2025 CAIDA snapshot, largest component).

Statistic	Value
Autonomous Systems (nodes)	78,144
Routing relationships (edges)	571,399
Mean connections per AS	14.62
Maximum connections	9,805
Graph density	0.000187
Detected communities (Louvain)	143
Largest community size	23,067

Plain English: Most ASes are small — a university, a regional ISP, a corporate network. A tiny fraction are enormous transit carriers that connect thousands of neighbors. The internet is not a flat mesh. It is a hierarchy wearing a mesh costume.

5 Methods (without the math wall)

Every Hidden Maps episode uses the same four ideas. Here is what they mean for the internet:

Idea	Internet translation
Hub (many connections)	A transit provider or hyperscaler with thousands of BGP neighbors
Chokepoint (high betweenness)	An AS on a large share of shortest paths — often a backbone, not a household name
Small-world	Regional clusters exist, but most ASes are only a few hops from most others
Random vs. targeted failure	Random: small ISPs vanish. Targeted: backbone ASes removed in order of centrality

Betweenness is **sampled** (800 random source nodes) because the full graph has tens of thousands of nodes — exact computation would take hours. Community detection uses the **Louvain**

algorithm: it finds clusters of ASes that route heavily among themselves, then reveals which ASes bridge those clusters.

6 Results

6.1 Unequal connections (the hub story)

Unable to display output for mime type(s): text/html

(a) How unequal is the internet? This chart shows the fraction of ASes that have at least k connections. The straight-line tail on log scales means a few giants dominate — classic heavy-tailed structure.

Unable to display output for mime type(s): text/html

(b)

Figure 1

The maximum-degree AS in this snapshot has 9,805 neighbors. The mean is 14.6. That gap — average two connections, maximum in the thousands — is the internet’s hub story in one sentence.

Unlike the power grid (engineered, truncated tails), the AS graph looks closer to the **preferential attachment** world: networks that grow big attract more connections because they are already big (Albert et al. 2000). This is why episodes 1 and 4 feel similar, while episode 2 (the grid) does not.

6.2 Small-world structure

Table 2: Small-world comparison. The ‘random graph’ row is a hypothetical internet where every AS connected to the same average number of neighbors but with no structure.

Table 2

Metric	Observed	Random graph	Plain English
Avg. clustering C	0.2914	0.000187	Neighbors of a network tend to connect to each other
Avg. path length L (sampled)	3.68	4.20	Few handoffs cross the internet on average
Small-world index	1775.6	—	Much more clustered than random, still short paths

The clustering coefficient is tiny in absolute terms — the internet is sparse — but **enormous relative to a random graph** with the same size and average degree. Translation: AS neighborhoods are not random wiring. Regional clusters (countries, hosting enclaves, academic networks) connect densely among themselves, while a backbone layer keeps global path lengths short (Watts and Strogatz 1998).

6.3 Who sits on the chokepoints?

How to read this chart: Brand fame and structural importance diverge. Google, Amazon, and Cloudflare are famous; transit carriers like Cogent, Lumen, NTT, and Hurricane Electric are the plumbing. When news reports say “a backbone provider had an outage,” these are the categories of players involved.

Unable to display output for mime type(s): text/html

Figure 2: Top 15 ASes by sampled betweenness centrality. These are structural bottlenecks — not necessarily the brands on your bill.

High betweenness does not mean evil or incompetent. It means **load-bearing under shortest-path routing** — the same compliment and curse that hub airports carry.

6.4 Communities: neighborhoods of the internet

Unable to display output for mime type(s): text/html

Figure 3: Distribution of community sizes detected by Louvain clustering. A few large neighborhoods contain many ASes; most communities are small.

Louvain finds 143 communities — routing neighborhoods that talk to themselves more than to outsiders. The largest contains 23,067 ASes. The internet is not one blob; it is continents of cooperation stitched by bridge ASes. Peering disputes hurt when they sever those bridges.

6.5 Robustness: random outages vs. hitting the backbones

Unable to display output for mime type(s): text/html

Figure 4: The series signature chart. Random AS removal (blue) vs. targeted removal of high-betweenness ASes (orange). The gap is the whole lesson.

This is the pattern across all four episodes:

- **Random failures** (small ISPs going offline, scattered outages) shrink the map slowly.
- **Targeted removal** of chokepoint ASes fractures global reachability fast.

The internet is **robust to ordinary noise, fragile to structural strikes** — not because it is poorly designed, but because shortest-path routing concentrates traffic on a subset of transit ASes. That concentration is efficient on good days and memorable on bad ones.

7 Discussion: What this does and does not claim

This map is a contract graph, not a traffic graph. Two ASes may be connected by BGP policy but carry little data between them. A weighted analysis using traffic matrices would sharpen operational conclusions. CAIDA’s relationships file is the standard structural starting point precisely because traffic data is proprietary.

BGP is not physics. Routing policy, filtering, and business disputes mean shortest paths in this graph are approximations of real packet paths. They are good approximations for structural questions — who *could* carry traffic — which is what robustness curves measure.

Sampling matters. Betweenness values here are estimates. Rank order of the top transit ASes is stable; precise numeric rankings below the top tier should be read as directional, not definitive.

For non-engineers, the takeaway is simple: The internet looks decentralized because no single company owns it. Topologically, it is highly centralized — a small set of ASes holds the map together. That is not a conspiracy; it is the economics of transit.

8 Conclusion

Episode 4 completes the arc from sky to grid to street to protocol. The vocabulary never changed:

- Dots and lines
- Hubs and chokepoints
- Random bad luck vs. the right failure
- Maps you never see until they break

The flight delay, the blackout, the detour, and the Cloudflare page were never separate stories. They were one story told by different graphs.

Next in Hidden Maps: biological and citation networks — connectomes, food webs, and the patent graphs that document how ideas borrow from ideas.

Part of the [Hidden Maps](#) series · Troy Altus · June 2026

Albert, Réka, Hawoong Jeong, and Albert-László Barabási. 2000. “Error and Attack Tolerance of Complex Networks.” *Nature* 406: 378–82. <https://doi.org/10.1038/35019019>.

CAIDA. 2025. *CAIDA AS Relationships Dataset, Serial-2*. https://catalog.caida.org/dataset/as_relationships_serial_2.

Luckie, Matthew, Bradley Huffaker, Amogh Dhamdhere, kc claffy, and Emile Aben. 2013. “AS Relationships, Customer Cones, and Validation.” *Proceedings of the 2013 Internet Measurement Conference*.

Watts, Duncan J, and Steven H Strogatz. 1998. “Collective Dynamics of ‘Small-World’ Networks.” *Nature* 393: 440–42. <https://doi.org/10.1038/30918>.