Gnatcatcher

Global Network Address Translation Combined with Audited and Trusted CDN or HTTP-proxy Eliminating Reidentification

Origins

- Privacy Sandbox aims to eliminate cross site identity linking

- Deprecating 3P cookies is a first step

- Need to address other mechanisms for identity linking, including IP

Ways to provide IP Address Privacy

- Give each person many addresses (e.g. one per tab)
- Give many people one address
- Make servers not use addresses for identity

Give each person many addresses

- Only possible with IPv6.
- With SLAAC, all addresses would still have same prefix, no actual privacy gains.
- Randomizing prefixes reduces hierarchical nature of Internet and makes packet routing slow and difficult at scale.

Gnatcatcher!

Hybrid of the two remaining solutions:

- Give many people one address: Near-path NAT
- Make servers not use addresses for identity: Willful IP Blindness

Willful IP Blindness

Servers willfully attest to not using IP addresses to track users, then volunteer to be audited to verify the degree to which they use IP addresses as identifiers.

Audit issues certificate site can show to browser to adjust things like a Privacy Budget accordingly.

Auditing: IP address use for packet routing

The hope of Willful IP Blindness is that there can be a delineation in serving stacks between:

- Serving stack layers that deal with IP addresses but not application data
- Application back-end that deals with URL and user ID but not IP addresses

A CDN could offer Willful IP Blindness conformance as a feature to the services they host.

Auditing: IP address use for server selection

- Sites need to select server based on proximity to user and content availability.
- Tricky because it can involve some application data (which resource being requested) and IP addresses (which server is closest).
- Have to fallback to auditing to ensure IP address not used for tracking.

Auditing: IP address use for anti-abuse

- See “Anti-abuse applications of IP” presentation for use case examples.
- Some anti-abuse needs can be fulfilled at the CDN level which makes auditing easier due to the layered nature of serving stacks.
- Anti-abuse mechanisms at deeper layers need to be audited to ensure they are compartmentalized and separated from other application back-ends.
- The hope is privacy-preserving alternatives (e.g. Trust Tokens) can decrease the reliance on IP addresses, thus making auditing easier.

Auditing: IP address use for geographic inference

- A site might need to know a user’s coarse geographic location to apply region-specific treatments as needed (e.g., GDPR, CCPA, etc).
- Willful IP Blindness will need to allow for this and audit use cases.

Auditing: IP address use for rare event investigation

In rare cases investigating site traffic using IP addresses is necessary.

- Debugging specific performance issues
- Investigating dangerous abuse

Auditing needs to ensure:

- “Break glass” access is required
- Access is logged in an auditable log with motivation provided
- Limited to a small fraction of all IP addresses/traffic

Willful IP Blindness & Privacy budget

- 7B people means 33 bits of entropy needed to be uniquely-identifying.
- IP addresses pose ~27 bits of entropy.
- 27 is almost all of 33 😟
- Doesn’t leave much budget for non-blind sites
- (Also points to the importance of solving for IP Privacy)

Gnatcatcher: Willful IP Blindness & Near-path NAT

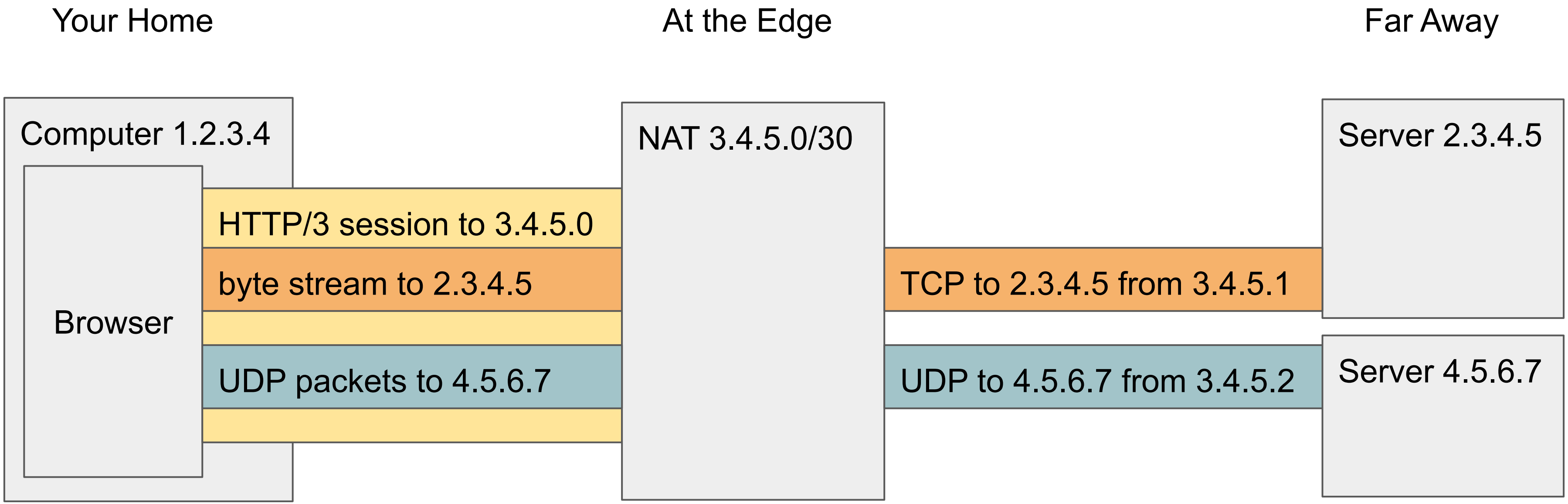
By including Near-path NAT, Gnatcatcher can be a complete solution, protecting users’ privacy across the web.

By including Willful IP Blindness, sites that need to perform cross-site anti-abuse measures can still continue to operate.

Near-path NAT: Idea

- NAT at the CDN level (at the edge)
- Individual NAT machine called an IP Privatizing Server (IPPS).
- Prevents a browser’s IP address from being used as a cross site identifier by:
    - IPPS’s IP addresses are shared amongst thousands of browsers.
    - IPPS’s IP,port tuple not stable across top-level domains.

Near-path NAT: Diagram



Near-path NAT: Performance

“Near-path”:

- By NATing at the edge (i.e. nearly on-path) the routing length isn’t increased.
- Avoids routing traffic to and from a datacenter that may be far off path.

Near-path NAT: Supporting Geo-IP

- By definition edge nodes are physically close to clients

- IP addresses will be useful for rough geolocation

- Sites can apply region-specific treatments as needed (e.g., GDPR, CCPA, etc.).

Near-path NAT: Best-effort legacy anti-abuse facilitated

Proposals like Trust Tokens can satisfy some anti-abuse use cases, lessening reliance on IP.

Near-path NAT facilitates first party anti-abuse:

➢ IPPSs should maintain a stable IP/port tuple per client and top-level-origin pair.

Within a site IP-address-based anti-abuse mechanisms can continue.

Where possible (i.e. IPv6) maintain stable & unique IP per client/top-level-origin.

Near-path NAT: Implementation exploration

IETF MASQUE WG working to extend HTTP/3 with the proxying capabilities that Near-path NAT needs.

MASQUE could proxy the browser’s TCP and UDP traffic streams through a single HTTP/3 connection to the IPPS.

MASQUE leverages HTTP/3:

- performance advantages of the modern HTTP/3 protocol
- powerful encryption of HTTP/3
- already deployed in browsers and servers