Held Hostage?

The Influence of Major ASes and CDNs on the Internet

Original Idea

- *The Internet is strongly hierarchical*

  - *Original maps (Rexford 2001) show the "Inner Core" lies in free-speech countries*

    - *US, France, Sweden*

- *But the Internet has changed a great deal.*

  - *How large is the inner core today?*
    How much lies in censorious countries?

First Round Approach

* *Based on : publicly available BGP data.*

    * *Routeviews Project*

* *Compute paths from every AS to "home AS" of target website. (Algorithm by Gao)*

* *Vary target websites and find common heavy hitters.*

First Round Results

- *"Core of the Internet" : 30 ASes*

  - *Loose term.*
    We mean, the heavy hitters that intercept >90% of paths to all target websites. (Alexa top-10, top-20 ...)

  - *Not a **true** backbone.*
    We can replace some of these with others in top-50 heavy hitters and still intercept >90% paths.

Ideas so Far - 1

- *The Internet has grown dramatically in 16 years (from 10,000 to 60,000 ASes) but the inner core has not (from 20 to 30 ASes).*

Ideas so Far - 11

- *Roughly one-third of the inner core is hostile.*
  (e.g. AS 4134, AS 4837 ... Great Firewall of China)

  - *Filtering by these ASes most likely affects transit traffic from downstream countries (collateral damage).*

  - *We **should** be worried about collateral damage from the censorship mechanisms in these ASes.*

Ideas so Far - III

- *Approx. 82% of the paths transit through core ASes in the United States.*

  - *Much greater than Russia (11%) or China (9%)*

- *But the US has given up net neutrality. (Dec 2017)*

  - *Perhaps **throttling** by US backbone providers will become a greater threat to open Internet access, than filtering by Russia, China, etc.?*

Problems - 1

- *Naive model of Internet Routing*

  - *Our model assumes that every site goes to the main server - e.g. google.com in Mountain View - and not to the closest local mirror.*

  - *In reality, much of the traffic is carried by CDNs (and not by AS-IXP routes).*

Problems - II

- *The AS relationships are well known (using Giotsas approach) - not all paths are valley free*

  - *But when stitching them together into paths, we still use Gao's algorithm ... assumes valley free paths*

  - *Needed : better approach to computing paths!*

- *Routeviews RIBs "biased toward big ASes" (Gregori)*

  - *Possible : rerun experiment with BGP tables from Isolario*

Going Forward

- *What paths do **actual** packets take? (including impact of CDN)*

- *How can we directly find impact of :*

  - *Filtering by censorious countries?*

  - *Throttling by ISPs in US?*

Importance of CDN

- *Take large sample of target websites*

  - *Alexa 10 k? (possibly 100 k?)*

- *From vantage points, see where the traffic is going when targeting these websites.*

  - *first cut :* dig <target website> *on vantage point*

  - *possible : confirm by running* traceroute

Importance of CDN

- *Common host serving many websites ... likely edge of CDN*

  - *confirm using reverse DNS* (dig -x) *and* whois

- *How many of these **real** paths are intercepted by censorious ASes?*

- *Also : direct measurement of impact of CDNs*

  - *Fraction of paths rerouted to CDN local cache*

  - *Dataset of savings in path length
    (compared vs. path to original server)*

Net Neutrality

- *Idea : try to identify targeted websites*

  - *torrent websites, music websites, porn*

  - *these are likely to be throttled by transit companies (which are also content companies)*

- *For each sensitive website, choose some peers*

  - *similar traffic rank, hosted in same AS*

Net Neutrality

- *From various vantage points, measure bandwidth to sensitive website AND to peers (using* abget)

  - *If sensitive website is throttled, it will be an outlier*

  - *Locate bottleneck (using* pathneck)
    Check to see if US ASes are doing the throttling

Net Neutrality

- *Particularly valuable as a longitudinal study*

  - *How the US became less free over time as a result of Net Neutrality repeal*

What do we want?

- *Vantage points!*

  - *To run* dig, traceroute

  - *To run* abget, pathneck

- *Comments and corrections.*

- *Better approach? Better tools?*