CAIDA, Tools

Miquel Garcia-Marron Coma
15/05/2023

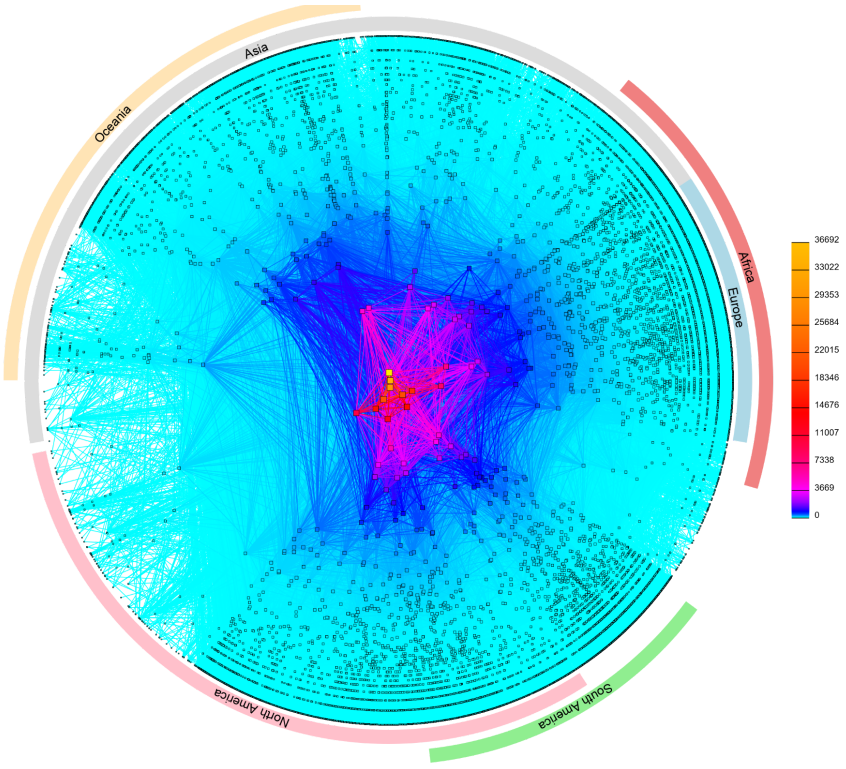
Index

1.Introducció

La Cooperativa per el Anàlisis de Dades d'Internet mes conegut coma a CAIDA es una organització sense ànim de lucre que va ser fundada l'any 1997. En aquests anys Internet estava en una fase d'evolució i creixement on les bases del que es el Internet actual no estaven creades. En resposta a aquestes demandes l'Universitat de California de San Diego va fundar aquesta entitat per investigar i intentar entendre el funcionament de Internet per a millorar el seu rendiment i funcionament.

Actualment CAIDA opera amb diverses eines i softwares que permeten portar una monitorització de la xarxa per a poder detectar fallades de seguretat o problemes amb la mateixa. També ofereixen serveis a empreses i col·laboren amb diverses entitats com poden ser NSF o el MIT que conjuntament amb experts del sector colaboren per identificar problemes.

En aquest treball indagarem en les eines més populars i utils que utilitza CAIDA actualment per el analisis del tràfic i detecció de problemes.



2.Eines

CAIDA utilitza varies eines per monitoritzar i fer estudis de xarxa. Aquestes les podem agrupar segons el propòsit pel qual han estat creades. Alguns d'aquests propòsits poden ser Analisis de Traffic, Mesurament de Topologia, Anàlisis de Encaminament, Segurtetat, Visualitzacio, …

A continuació veurem algunes eines que cobreixen alguns dels propòsits anteriors.

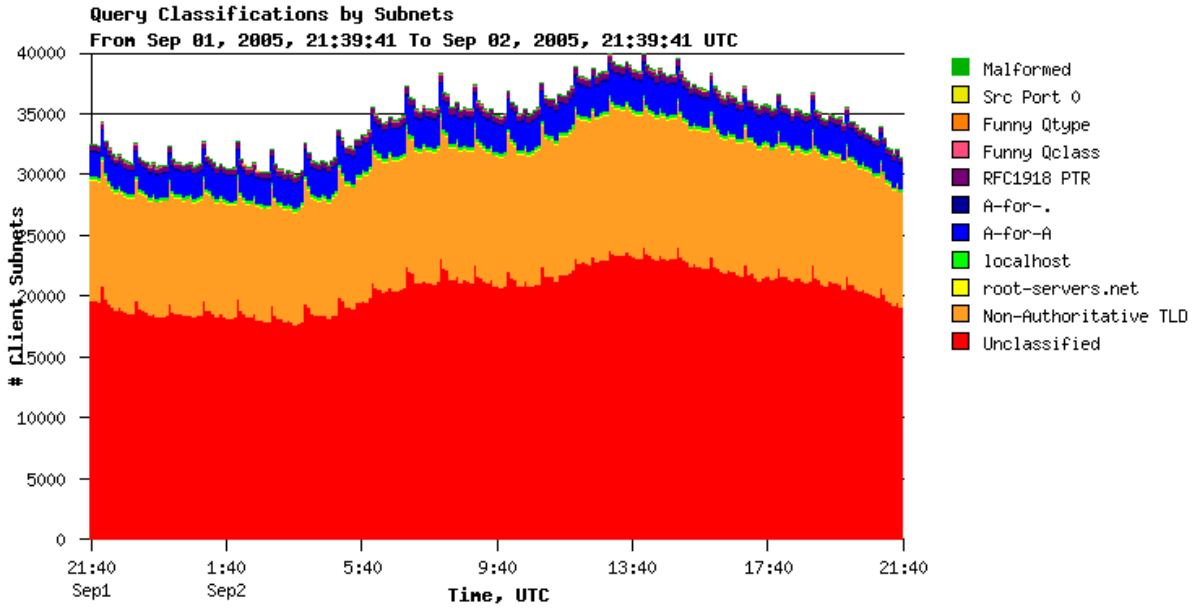
DSC (DNS Statistics Collector)

DSC es un sistema de recoleccio i exploracio de estadístiques de DNS amb molta demanda. Aquesta eina esta formada per dos grans parts:

La primera consisteix en un apartat de coleccio de dades. Aquesta es du a terme dins del mateix servidor DNS o sino es possible dins de la mateixa xarxa amb el us de un switch amb port mirroring de manera que tots el paquets queadarien replicats en el nostre port.

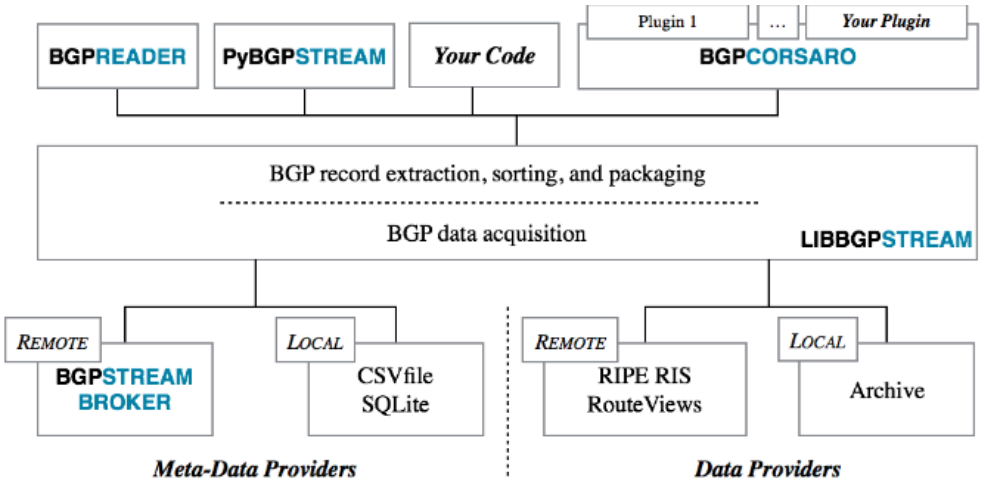
La segona part gestiona les dades aportades per la primera part, aquesta s'encarrega de crear gràfiques i taules per fer-ne un ús posterior d'estudi.

Un exemple n'és el gràfic següent, processat amb aquesta eina ens mostra les requests de clients de un DNS classificat per subnets:



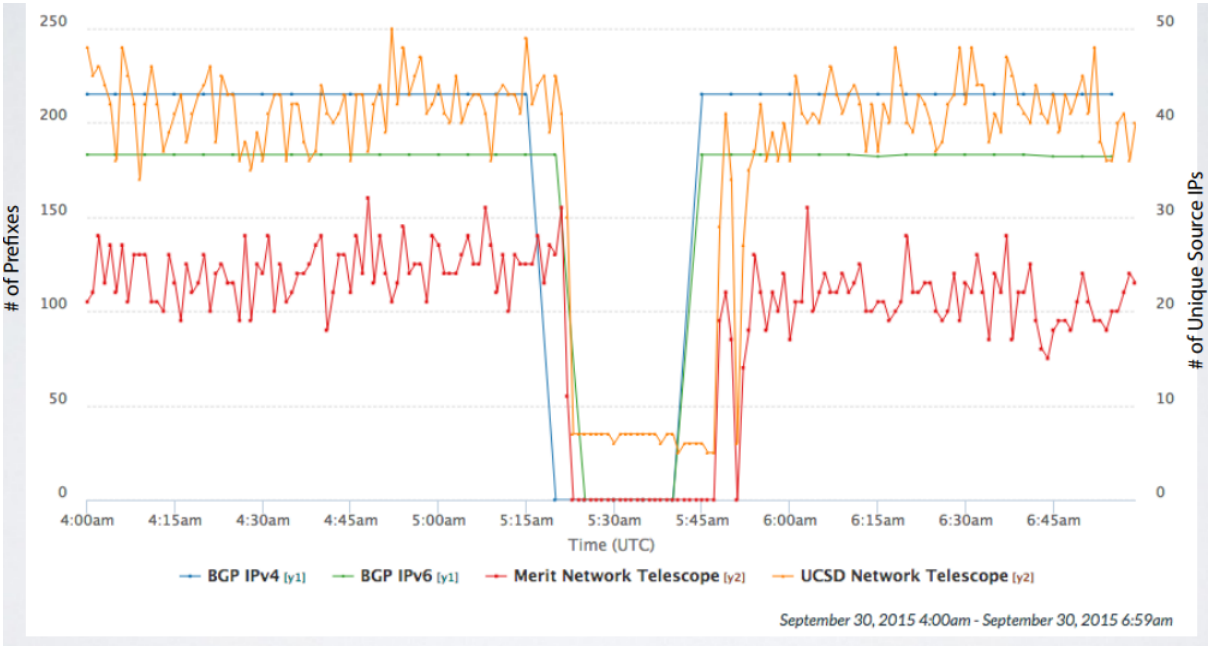
BGPStream

BGPStream es una eina open source desarrollada per CAIDA que s'usa per el anàlisis del protocol BGP. El protocol BGP es un protocol que s'usa per enrutar sistemes autònoms que determinen una part molt gran de la xarxa.



Com podem veure en la imatge anterior com funciona BGP Stream. A grans trets compten amb dos grans fonts d'informació, la primera es una base de dades que disposa de dades no a temps real sino mesurades amb anterioritat. Aquestesson usades per a fer comparacions temporals. I l'altre font d'informació la conformen AS que usen BGP per enrutar paquets. Aquertes ofereixen informació a BGPStream, aquesta informació és a temps real i permet determinar el estat actual de la xarxa. A part de que les AS ofereixin el seu traffic també tenim a institucions com RIPE NCC que té actualment 27 col·lectors recollint dades a temps real. Una vegada es tenen aquestes dades els investigadors mitjançant les llibreries que ofereix aquest software pot estudiar la topologia de la xarxa actual i la seva salut.

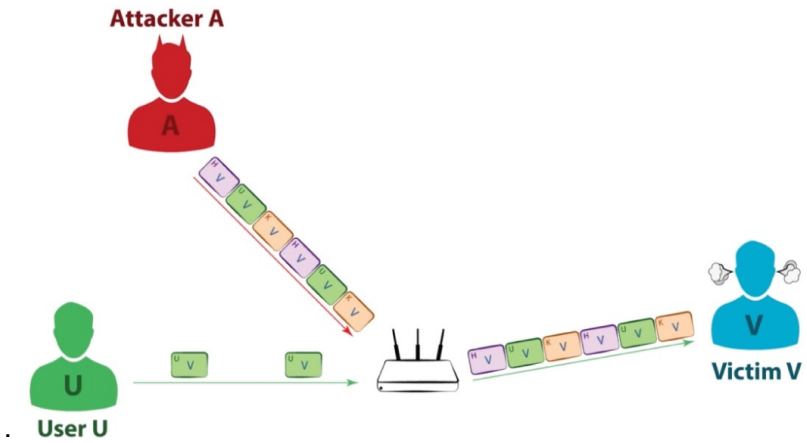
A continuació es pot veure un exemple de analisis de xarxa per BGP en aquest cas es tracta de una grafica que mostra el tràfic de la xarxa el 30 de Setembre de 2015 justament quan va caure la xarxa de Time Warner Cable LLC amb numero de sistema autonom AS11351:



Veiem que la caiguda es notable desde les 5 de la matinada fins les 6.

Spoofer

Spoofer es una eina que controla el nivell de spoofing que hi ha a Internet. El spoofing es una técnica que es utilitzada pels cibercriminals per suplantar la identitat d'un individu amb fins delictius.



A grans trets en tenim de 5 tipus:

- Spoofing amb Emails: Es una de les tècniques mes usades actualment, el atacant envia una gran quantitat de emails a un munt de víctimes fent-se passar per una empresa o entitat bancaria sol·licitant credencials.

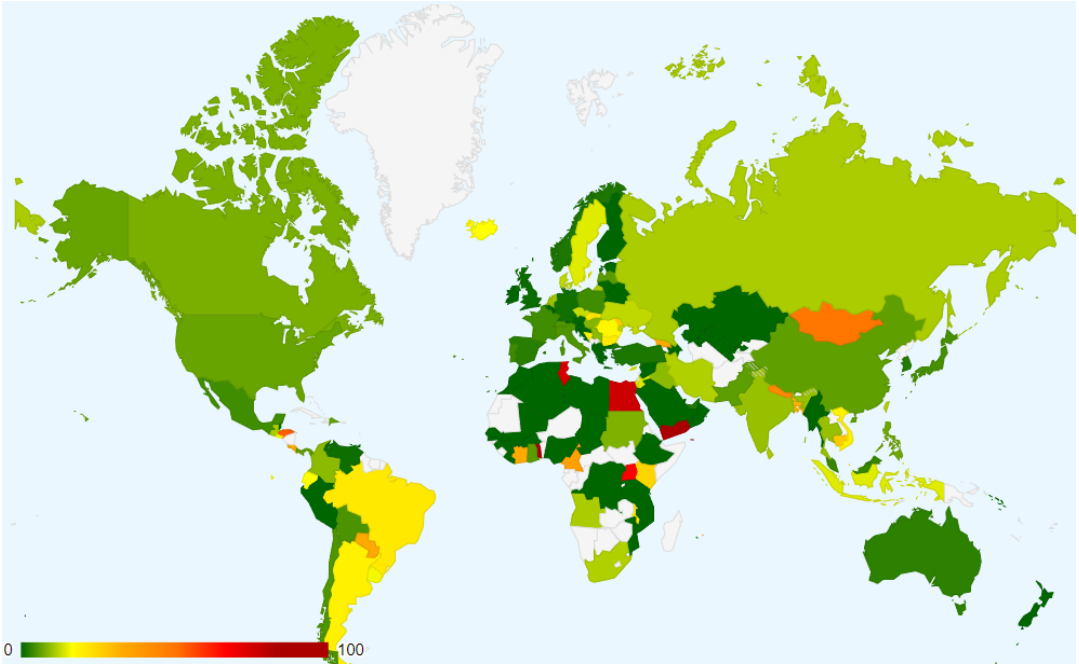
- Spoofing de DNS: En aquest tipus de spoofing el hacker canvia entrades del DNS per aconseguir que la víctima al accedir a una página web determinada el DNS al resoldre la peticio el redirigeixi a pàgines web malicioses.

- Spoofing de IP: El atacant suplanta la seva IP amb una altra de falsa es un procediment semblant al que usen les VPN.

- Spoofing de DDoS: Es un tipus de spoofing que s'usa per a fer atacs de Denegació de Serveis, els atacants busquen ordinadors vulnerables per a poder instalar los softwares maliciosos per acabar creant botnets que les usara per realitzar moltes peticions a serves els quals vulgui denegar.

- Spoofing ARP: El atacant es col·loca en una xarxa local i poden suplantar l'adreça MAC de la víctima llavors aquest es posara enmig de una comunicació i podra veure el tràfic de la connexió.

Aquestes no son totes les maneres de fer spoofing, existeixen moltes altres tecniques. Spoofer resol aquests problemes amb l'ajuda de firmes digitals i també a la comprovació de l'origen dels paquets evita tot aquest tipus de spoofing. Apart de evitar el spoofing també fa estudis detallats d'aquests atacs. A continuació podem veure una gráfica que indica el percentatge de blocs de IP que mostren sintomes de spoofing. Com més vermell mes indicis de spoofing.

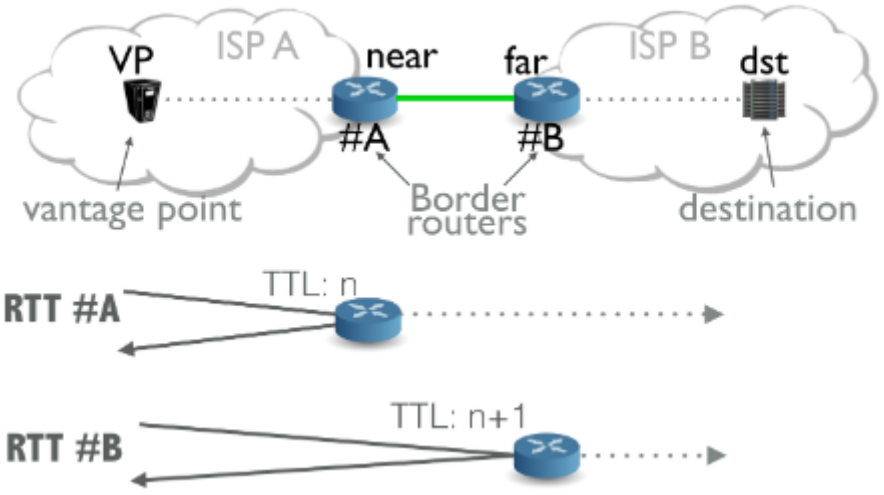


Es pot veure que la gran majoria de xarxes amb possibilitat de spoofing son països del tercer mon.

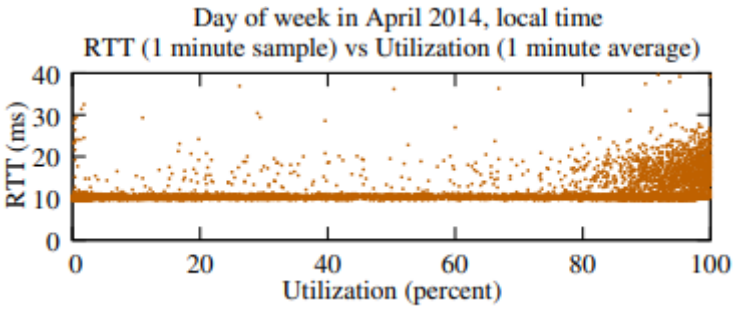
MANIC (Measurement and ANalysis of Internet Congestion)

Aquest eina es usada per CAIDA per mesurar els nivells de congestió de la xarxa. Aquesta eina mesura la quantitat de congestió a nivell interdomain o sigui la congestió donada entre AS no dins de cada una.

Aquesta eina utilitza un mètode que s'anomena TLSP (Time-Sequence Latency Probing). Es comença per seleccionar els punt on volem medir la congestió. En el nostre cas seran ISP A i el ISP B, fixem dos punts, en el nostre cas seran els routers de sortida de cada AS. El punt near sera el router de sortida de el ISP A origen i el punt fares el router de arribada del ISP B destí. També tindrem un punt de recol·lecció de dades que es diu Vantatge Point aquest és administrat per CAIDA. Una vegada fixats els punt enviarem missatges desde el Vantatge Point a near amb un TTL n on n es els salts que ha de fer el missatge per arribar a near, llavors també farem el mateix per B en aquest cas el TTL sera n + 1 o sigui el numero de salts per arribar a near mes un. Llavors fent la resta del RTT entre near i far podem veure quan es triga a passar per l'enllaç en cas que estigues congestionat el RTT de far seria molt més gran que el de near i si el comparessim amb resultats anteriors podem veure quan de congestionat esta.



Una vegada recollits els resultats es fan gràfiques i estudis per determinar la congestió de la xarxa. A continuació tenim una gráfica generada amb aquesta eina on podem veure el RTT del enllaç envers la utilització del mateix. Podem veure que al traspassar el llindar del 80% d'utilització aquest es comença a saturar i les marques de RTT comencen a augmentar.



MIDAR (Monotonic ID-based Alias Resolution)

Midar es un software destinat a ajudar als creadors de topologies de mapes. Quan parlem de la creació de mapes de topologies de Internet hem de tindre en compte quines IP pertanyen al metix router o sigui dit d'una altre manera es requereix de l'agrupació de totes les IP de un router (per interfices) aquest procés es diu IP alias resolution i aquesta eina es la encarregada de dur-ho a terme.

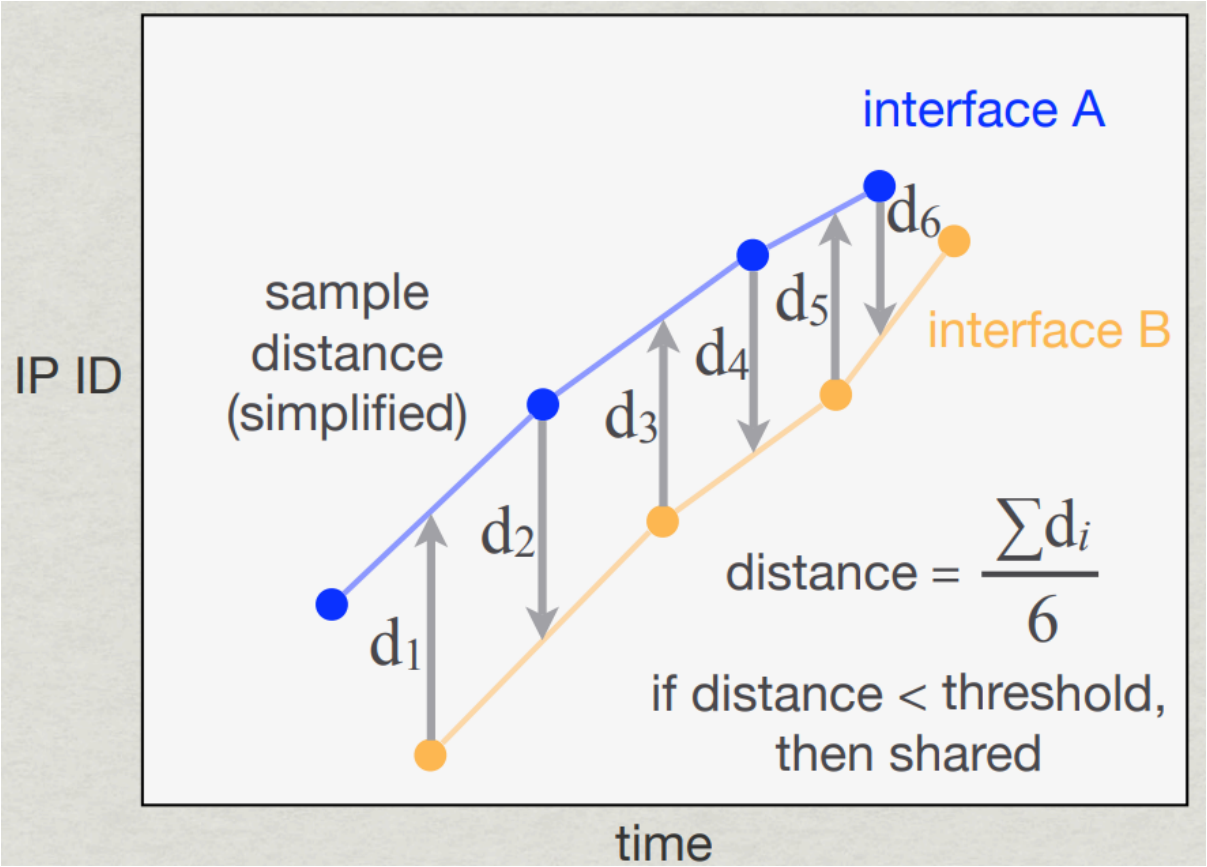
Per a dur a terme aquest procés es duu a terme en 4 passos:

- Estimació: per a cada adreça, determinem la velocitat d'enviament de paquets i el millor mètode de sonda per utilitzar-lo en etapes posteriors. per a cada adreça objectiu es determina la velocitat a llançar els missatges mes idonia i també el mètode més eficient..

- Descobriment: es sonden les adreces objectiu en busca de posibles adreçes parelles a la inicial..

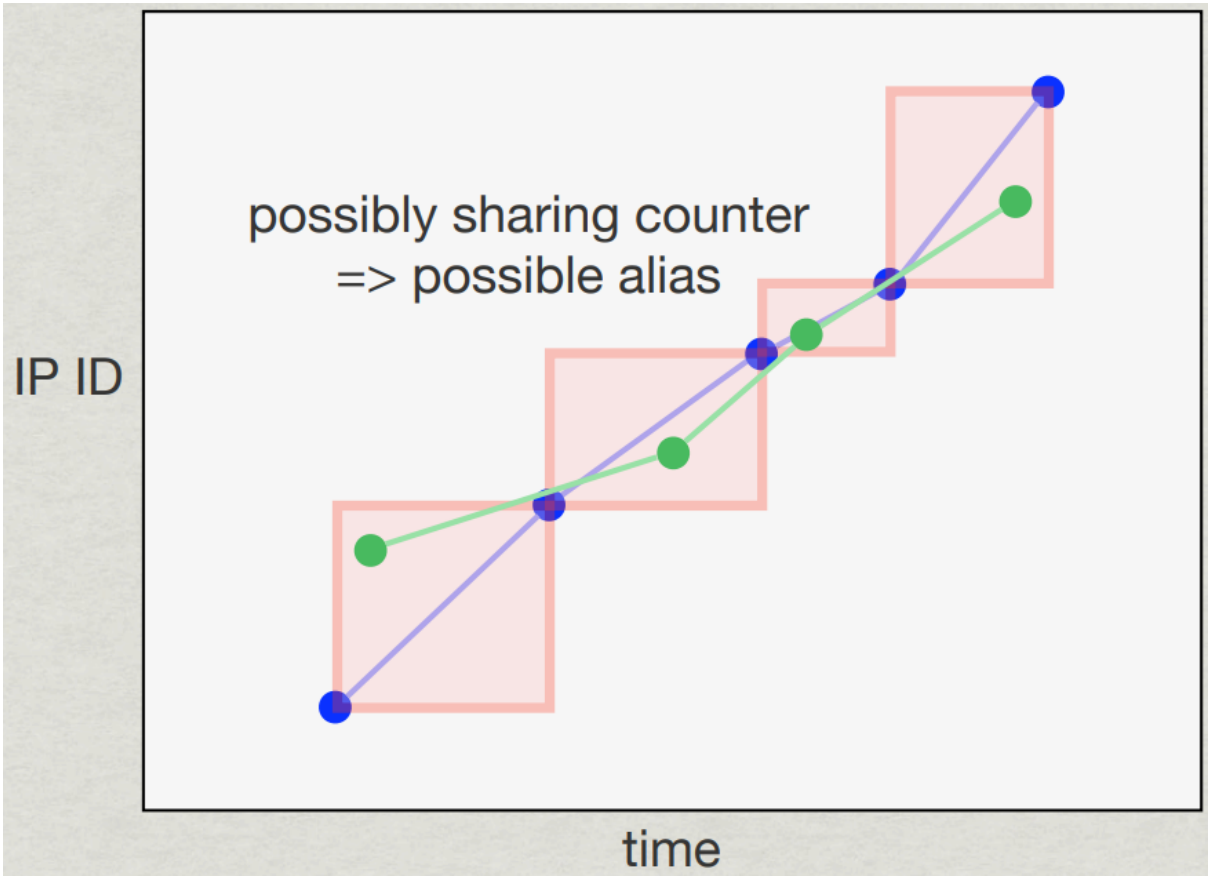
- Corroboració i Eliminació: es fa un resondatge unes quantes vegades per a assegurar-se que no agafem cap fals positiu com a vàlid.

Quan parlem de sondatge hi han moltes tècniques aqui explicaré només el mètode RadarGun. Aquest métode utiliza el camp IP ID del paquet TCP que normalment és utilitzat en la fragmentació. Al enviar paquets a un encaminador aquests els retornara amb un IP ID diferent, aquest IP ID en realitat no serà tant diferent ja que el encaminador utilitza un comptador incremental de manera que paquets que van a interficies diferents tindran el camp IP ID forsa semblant, de fet si anem llençant paquets aquests seguiran una funció semblant a la de la gràfica següent per a cada interfície:



Resumidament si calculem la mitjana de distàncies entre les dos funcions i es menor a un llindar que determina si les interfaces pertanyen al mateix encaminador llavors voldrà dir que les interfaces son del mateix encaminador. Aquest mètode presenta fallades ja que mai sabrem quin es el threshold òptim per tant sempre sortiran falsos positius i d'altre banda tindrem que aquest mètode requereix de molts recursos als encaminadors.

Per a resoldre aquest problema MIDAR ofereix un altre mètode que s´anomena Monotonic Bounds, que es un mètode que funciona amb finestres llavors no abarca tants recursos i funciona igual que RadarGun pero en mes de comparar la mitjana amb un llindar prefixat. A continuació una gráfica que representa el funcionament.



Es creen unes areas marcades per cada marca tempora per una interfas A i es comprova si cada punt de la interfas B es dins l'àrea de manera que aquest mètode és més eficient i consumeix menys recursos ja que funciona per finestres temporals.

3. Conclusions

Hem vist algunes de les moltes eines usades per al tractament i anàlisis de dades que utilitza CAIDA. Hem vist desde análisis de DNS, control de BGP per mesurar la Topologia de internet, hem vist també alguna eina de seguretat com pot ser Spoofer per evitar els atacs de spoofing. D'altra banda també hem vist eines de suport de la visualització com pot ser MIDAR que ajuda a detectar les IP que provenen del mateix router per suportar la visualització de gràfiques. I també per descomptat MANIAC que ajudava a detectar quan tenim congestió a la nostra xarxa per a poder evitar-ho en properes vegades.

Totes aquestes eines donen suport a CAIDA pero no son les úniques en aquest moment CAIDA te un total de 43 softwares per el análisis. Molts d'aquests son similars als explicats en aquest treball, les diferències les tenim en les maneres de recollida i tractament de dades. D'altra banda també hi han altres eines que no hem vist i també tenen molta importància.

Per acabar podem dir que CAIDA té moltes eines per analitzar la xarxa i la gran majoria no les hem vist en aquest treball, pero totes serveixen a propòsits molt importants que ajuden a l'organització a l'analisis, detecció i prevenció de futurs problemes a Internet.

4. Referències

- *CAIDA*. (s/f). CAIDA. Recuperado el 15 de mayo de 2023, de https://www.caida.org/
- *Dsc: A DNS Statistics Collector*. (s/f). Measurement-factory.com. Recuperado el 15 de mayo de 2023, de http://dns.measurement-factory.com/tools/dsc/
- Keys, K., Hyun, Y., & Luckie, M. (2010). *Internet-Scale Alias Resolution with MIDAR*. Caida.org. https://www.caida.org/workshops/isma/1002/slides/aims1002_yhyun_midar.pdf
- King, A. (s/f). *BGPStream*. Caida.org. Recuperado el 15 de mayo de 2023, de https://bgpstream.caida.org/docs/api
- *Manic API doc. With swagger UI*. (s/f). Caida.org. Recuperado el 15 de mayo de 2023, de https://api.manic.caida.org/v1/
- *MIDAR*. (2010, marzo 5). CAIDA. https://www.caida.org/catalog/software/midar/
- Orsini, C., King, A., Giordano, D., Giotsas, V., & Dainotti, A. (2016). BGPStream: A software framework for live and historical BGP data analysis. *Proceedings of the 2016 Internet Measurement Conference*.
- *Spoofer*. (s/f). CAIDA. Recuperado el 15 de mayo de 2023, de https://www.caida.org/projects/spoofer/

M'HE QUEDAT A LA CORRECIO A LA PART DE MANIC